

## MODULE 6- LECTURE 1

### TRANSGENIC SCIENCE IN PLANT IMPROVEMENT

#### 6-1.1 Introduction

Transgenic plants are developed by transferring or modifying genes from another organism by a diverse technique like physical, chemical and biological methods. Transgenic plants are obtained by introducing a gene into its genome with the help of vectors in order to develop a plant with new characteristics. This process of recombinant DNA technology is used for developing genetically modified plants to embolden the variety against pests, diseases and stress or to improve the quality of the product. These can also be achieved by conventional breeding if the genes for the desired characters are present in close or wild relatives else transgenesis is the only viable alternative. Apart from this, genetic engineering can overcome the other limitations of conventional breeding such as large space requirement, time consuming, and uncertainty of the result. It is not mandatory to opt for transgenics if the problem can be addressed by conventional breeding. Rational decision making can be done by following figure 6-1.2.1.

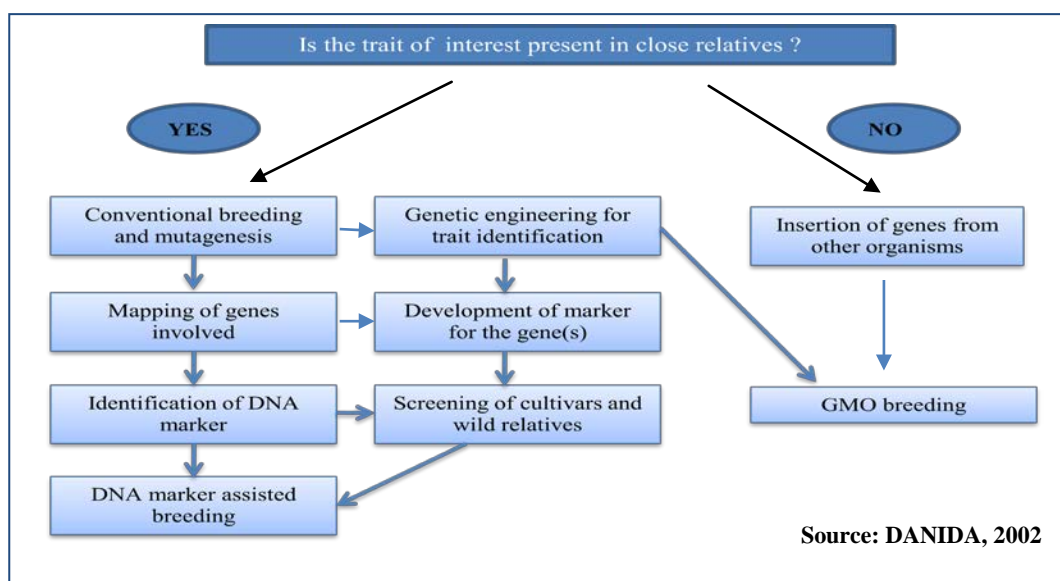


Figure 6-1.2.1 Decision making box by using both conventional and modern biotechnology approaches for crop breeding

## 6-1.2 Modern Plant Breeding

Modern plant breeding uses techniques of molecular biology to select the variations. This approach which combines biotechnology and molecular biology is known as molecular breeding.

### 6-1.2 .1 Steps Involved In Plant Breeding

Following steps are involved in both traditional as well as modern plant breeding-

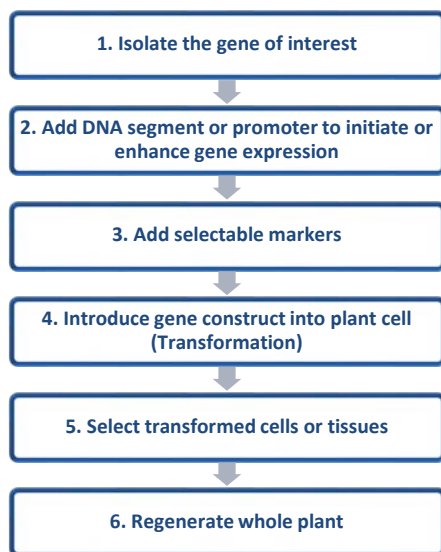
- 1. Generation (or identification) of variant** may be done simply by collection of wild type varieties or traditional varieties from farmers or by hybridization (crossing 2 or more plants species) and by induced mutation and induced polyploidy
- 2. Selection for desired characteristics** like grains for large seed size, seed dormancy, and non-shattering seed heads. It is possible now to select variants based on molecular markers.

These two steps are still followed for plant breeding but biotechnology has changed the breeding techniques. In this chapter we are going to focus more on the new technology of introducing characters to the plant called **Transgenic Technology**.

6-1.2 .2 Transgenic Technology results into genetic variation across kingdoms, rather than within a species or genus. Gene transfer is more precise than previous methods.

#### 6-1.2.2.1 Steps Involved In Production of Transgenic Plants

There are several steps involved in production of transgenic plants which is given in figure 6-1.2.2.1



**Figure 6-1.2.2.1 Steps Involved in production of transgenic plants**

#### 6-1.2.2.2 Isolation and Cloning the Gene of Interest

Identification and cloning of the gene of interest is a first **limiting step** in the transgenic development process. Locating, identifying, characterizing, and cloning genes of agricultural importance requires, a huge effort both in terms of human and financial capital. One of the earliest development is the introduction of insect resistance by transgenic technology. The discovery of **Bt Genes** has revolutionized plant transgenics.

Spores of the soil bacterium *Bacillus thuringiensis* (Bt) contain a **crystal (cry) protein ( $\delta$ -endotoxin)**. Inside insect gut, the crystals break apart and release a toxin that binds to and creates pores in the intestinal lining. Instead of whole gene **truncated cry gene** is used in **Bt crops**. Figure 6-1.2.2.2 shows the truncated cry gene structure.

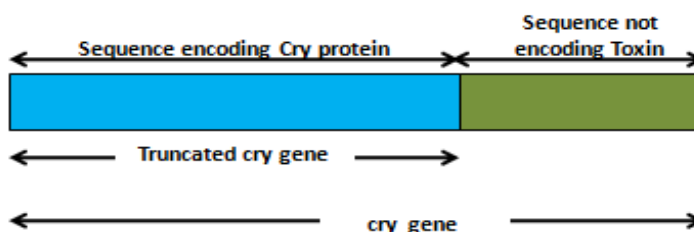


Figure 6-1.2.2.2 Gene sequence showing the truncated cry gene

#### Mechanism of Toxicity:

- *Bt* gene (also known as *cry* gene) was found in a gram positive bacteria *Bacillus thuringiensis*. The structure of *Bt* gene is shown in Figure 6-1.2.2.2
- This *Bt* gene is used in the production of insect resistant crops (genetically modified) and biological insecticides as well.
- *Bacillus thuringiensis*, during sporulation produces a toxic protein which possesses insecticidal activity against Lepidoptera, Coleoptera, Hymenoptera, Diptera and Nematode.
- This crystal protein is called as Cry protein, encoded by *cry* gene present on the plasmid (non-chromosomal gene).
- As soon as the Cry protein crystals reach the digestive tract of the insect, the prevailing alkaline condition there causes denaturation of the insoluble crystals.

- This denaturation makes the crystal soluble and prone to proteases activity in the gut of insect.
- Proteolysis of Cry crystal leads to release of cry toxin, which forms pore in the cell membrane of the gut by inserting themselves into it.
- The pore causes cell lysis and ultimately death of insects.

#### 6-1.2.2.1.2 Control of Gene Expression

- The level of gene expression is determined by regulatory sequences such as promoters as well as 5' UTR elements (Described in detail in module 5- lecture 5).
- **Transgene Promoters:** Most commonly used is the **CaMV 35S promoter** of cauliflower mosaic virus. It is a **constitutive** promoter (turned on all the time in all tissues), that gives high levels of expression in plants.
- Most commonly used terminator sequence is the nopaline synthase (*nos*) gene from *Agrobacterium tumefaciens*.

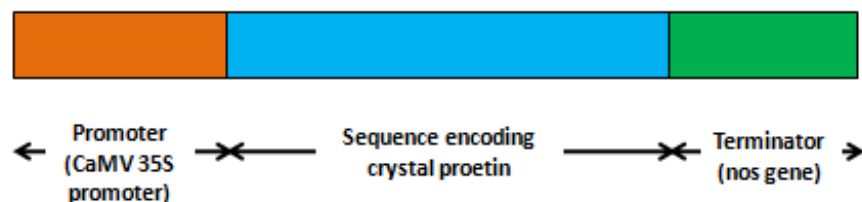


Figure 6-1.2 .2.1 .2 shows expression cassette of Bt gene along with promoter and terminator.

#### 6-1.2 .2.1.3 Selectable Markers

- Various selectable markers are used for the selection of transgenic plants (described in detail in Lecture 4 and 5 of Module 5).

#### 6-1.2.2.1.4 Plant Tissues Used For Transformation

The tissue must be capable of generating **callus** (undifferentiated tissue), from which the complete plant can be produced. The choice of tissue depends on the species. Some commonly used tissues are immature embryos, leaf disks, and apical meristems.

#### 6-1.2.2.1.5. Introduction of Gene Construct into Plant Cells (Transformation)

(We have discussed Transformation in Lecture 4&5 of Module 5)

The basic methods and techniques used for plant cell transformation are as follows:

- i) *Agrobacterium* mediated transformation
- ii) Agro-infection
- iii) Chloroplast transformation
- iv) Indirect gene transfer
- v) Electroporation
- vi) Biolistic (Gene gun) Method
- vii) Chemical method of gene transfer
- viii) Microinjection
- ix) Pollen Transformation
- x) Direct DNA uptake by mature zygotic embryos

#### 6-1.2.2.1.6. Selection of Transformants

- Cells/tissues in which new genes are incorporated into plant's DNA are identified and then grown in media containing antibiotics or herbicides.
- So during transformation, selectable marker gene (antibiotic resistant etc.) is incorporated such that, cells which are devoid of this plasmids get killed or their growth is arrested. Kanamycin, an antibiotic capable of killing plant cells.
- Since transformed plants contain kanamycin resistance gene, they only can survive.
- In the *Agrobacterium* infiltration method, the seed of the infiltrated plants are plated on agar containing kanamycin – the plant seed containing the introduced DNA will germinate (less in number) and will grow on agar plates containing Kanamycin.

#### 6-1.2.2.1.7. Regeneration of Transformant

- During regeneration, whole plants with inserted genes are developed through tissue culture.
- Cultured transformed plant cells are regenerated under appropriate conditions.
- Auxin/cytokinin ratio is important for plant regeneration.

- In the culture media high concentration of auxin (plant growth regulator), causes rooting of the cultured transformed cells. Further, shooting is initiated by increasing the concentration of cytokinin (plant growth regulator-phytohormones) and new plantlet develops from the culture.
- After that plantlets are transferred to soil for proper anatomical and physiological developments, and for better acclimatization to environment.

### **6-1.3 Confirmation of transformed plants**

The presence and activity of introduced gene is confirmed by observation of phenotype and by advanced methods as listed below.

- 1) **Southern blot:** In Southern blotting separated DNA fragments obtained after electrophoresis, are transferred to a filter membrane and subsequent fragment detection is accompanied by probe hybridization.
- 2) **Northern blot:** It is used to study gene expression by detection of mRNA (or isolated mRNA) in a plant sample.
- 3) **Western blot:** It is used to study the gene expression by detection of the protein produced by the transformed regenerated plants.

Details of these blotting techniques are explained in earlier chapter (Module 3 Lecture 4).

### **6-1.4 Evaluation of Transformed Plants**

- Plant is evaluated for the presence and activity of introduced gene.
- The effect of various environmental factors on the transgenic plant is also evaluated.
- Evaluation for food or feed safety.
- Evaluation of basic containment level is also very important.

### **6-1.5 Chloroplast Transformation**

The basic techniques we have discussed in lecture 5 of module 5. Here we are going to see some of the application of chloroplast engineering.

#### **6-1.5.1. Improvement of Plant Traits**

Plastid genome can be used to engineer many agronomic traits such as herbicide resistance, insect resistance, and antibiotic resistance.

### 6-1.5.2. Biotic Stress

Expression of insect resistance genes such as *cry* genes can be well performed in the plastid genome without modifications of codon usage or other sequence manipulations.

### 6-1.5.3. Abiotic Stress

Chloroplast engineering can be executed for the successful development of plants suffering from abiotic stresses like salinity, high or low temperature, drought etc. Some examples are listed below-

- The over-expression of enzymes required for Glycine betaine (GlyBet) biosynthesis in transgenic plants improve tolerance to various abiotic stresses.
- Chloroplast transformation has been used to transfer choline monooxygenase (BvCMO), an enzyme that catalyzes the conversion of choline into betaine aldehyde from beet (*Beta vulgaris*) into the plastid genome of tobacco.
- Transplastomic carrot plants expressing betaine aldehyde dehydrogenase (BADH) gene have also been developed using chloroplast engineering showing highest level of salt tolerance.

### 6-1.5.4. Production of biopharmaceuticals

Chloroplast engineering can also be employed for the production of biopharmaceutical proteins, antigens listed in Table 6-1.5.4.1. and 6-1.5.4.2.



Table 6-1.5.4.1 Depicts biopharmaceutical proteins expressed via chloroplast transformation

<b>Biopharmaceutical Proteins</b>	<b>Gene</b>	<b>Site of Integration</b>	<b>Promoter</b>	<b>5'/3' Regulatory Elements</b>
<b>Elastin Derived Polymer</b>	<i>EG 121</i>	<i>tm/trnA</i>	<i>Prrn</i>	<i>ggagg/TpsbA</i>
<b>Antimicrobial Peptide</b>	<i>MSI-99</i>	<i>tm/trnA</i>	<i>Prrn</i>	<i>ggagg/TpsbA</i>
<b>Insulin-like growth factor</b>	<i>IGF-1</i>	<i>tm/trnA</i>	<i>Prrn</i>	<i>PpsbA/TpsbA</i>
<b>Interferon alpha 5</b>	<i>INFα5</i>	<i>tm/trnA</i>	<i>Prrn</i>	<i>PpsbA/TpsbA</i>
<b>Interferon alpha 2b</b>	<i>INFα2b</i>	<i>tm/trnA</i>	<i>Prrn</i>	<i>PpsbA/TpsbA</i>
<b>Human serum Albumin</b>	<i>hsa</i>	<i>tm/trnA</i>	<i>Prrn</i> , <i>PpabA</i>	<i>ggagg<sup>a</sup>TpsbA /TpsbA</i>
<b>Monoclonal Antibodies</b>	<i>Guy's 13</i>	<i>tm/trnA</i>	<i>Prrn</i>	<i>ggagg/TpsbA</i>
<b>Interferon gamma</b>	<i>IFN-g</i>	<i>rbcL/accD</i>	<i>PpsbA</i>	<i>PpsbA/TpsbA</i>
<b>Human somatotropin</b>	<i>hST</i>	<i>trnV/rps 12/7</i>	<i>Prrn</i> , <i>PpabA</i>	<i>T7gene 10psbA' trps16</i>

Table 6-1.5.4.2 Depicts vaccine antigens expressed via chloroplast transformation

<b>Vaccine Antigens</b>	<b>Gene</b>	<b>Site of Integration</b>	<b>Promoter</b>	<b>5'/3' Regulatory Elements</b>
<b>Cholera toxin</b>	<i>ctxB</i>	<i>trnI/trnA</i>	<i>Prrn</i>	<i>ggagg/ TpsbA</i>
<b>Canine Parvovirus (CPV)</b>	<i>ctxB-2L21 gfp-2L21</i>	<i>trnI/trnA</i>	<i>Prrn</i>	<i>PpsbA/ TpsbA</i>
<b>Anthrax Protective antigen</b>	<i>pag</i>	<i>trnI/trnA</i>	<i>Prrn</i>	<i>PpsbA/ TpsbA</i>
<b>Plague Vaccine</b>	<i>caF1~LcrV</i>	<i>trnI/trnA</i>	<i>Prrn</i>	<i>PpsbA/ TpsbA</i>
<b>Tetanus toxin</b>	<i>tetC</i> (bacterial and synthetic)	<i>trnV/rps 12/7</i>	<i>Prrn</i>	<i>T7 gene 10, atpB/Trbcl</i>

### 6-1.5.5. Metabolic Pathway Engineering

- Chloroplast engineering represents an attractive alternative to conventional nuclear transgene expression for metabolic engineering. The strong transgene containment and possibility to stack multiple transgenes by linking them in operon makes this process widely exploited in chloroplast genome.

Synthesis of polyhydroxybutyrate (PHB), a bioplastic has been achieved by introduction of the most complex metabolic pathway into the chloroplast genome so far. Three enzymes of PHB biosynthesis are co-transcribed into the tobacco plastid genome. Significant accumulation of PHB in chloroplasts causes male sterility and severe growth retardation. The application of plastid transformation to metabolic pathway engineering is still restricted to few model species like agrobacterium.

- Metabolic engineering has been emerged as a promising technology in food crops. Recently, plastid transformation has been performed in tomato to alter carotenoid biosynthesis for the production of fruits with elevated contents of provitamin A ( $\beta$ -carotene), an important antioxidant and essential vitamin for human nutrition.

### 6-1.6 Examples of GM (Genetically modified) Crops or transgenic plants

Table 6-1.6 Crops from which transgenes might introgress into wild relatives

Crop	Main species	Progenitor species	Genomic structure, chromosome number (2n)	Transgenic varieties	Main wild relatives (2n) (distribution)	References*
Sorghum	<i>Sorghum bicolor</i>	<i>S. bicolor</i>	Two subgenomes of five chromosomes (2n=10)	None	<i>S. bicolor</i> (2n=20) (Old and New World) <i>Sorghum halepense</i> (2n=40) (Old and New World) <i>Sorghum propinquum</i> <i>Sorghum alatum</i>	106,107 65,66 66 66
Sunflower	<i>Helianthus annuus</i>	<i>H. annuus</i>	(2n=34)	Herbicide and insect tolerance in development	<i>H. annuus</i> (2n=34) (New World) <i>Helianthus petiolaris</i> (2n=34) (New World)	62,63 16,64
Canola	<i>Brassica napus</i>	<i>Brassica rapa</i> and <i>Brassica oleracea</i>	AACC (2n=38)	Herbicide tolerance commercially released, insect and disease stress tolerance	<i>B. rapa</i> (AA, 2n=20) (widespread in most canola growing areas)	57 <sup>†</sup>
	<i>B. rapa</i>	<i>B. rapa</i>	AA (2n=20)	Commercial release withdrawn	<i>B. rapa</i> (AA, 2n=20) (widespread in most canola growing areas)	–
Wheat	<i>Triticum aestivum</i>	Wild <i>Triticum</i> and <i>Aegilops</i> species	AABBDD (2n=42)	Herbicide tolerance developed, not yet commercially released	<i>Aegilops cylindrica</i> (DD, 2n=28) (North America)	54, 55
Sugar beet	<i>B. vulgaris</i> ssp. <i>vulgaris</i>	<i>B. vulgaris</i> ssp. <i>maritima</i>	(2n=18/27)	Herbicide and disease (virus) tolerance commercially released	<i>B. vulgaris</i> ssp. <i>maritima</i> (2n=18) (Old World)	52,53,108,109
Alfalfa	<i>Medicago sativa</i> ssp. <i>sativa</i>	<i>M. sativa</i>	(2n=32)	Herbicide and stress tolerance developed but not commercially released	<i>M. sativa</i> ssp. <i>sativa</i> (ssp. <i>caerulea</i> ) (2n=16) (Old and New World) <i>M. sativa</i> ssp. <i>falcata</i> (2n=16/32) (Old and New World)	50,51

(Source: C. Neal Stewart, Jr, Matthew D. Halfhill & Suzanne I. Warwick., ( 2003), *Transgene introgression from genetically modified crops to their wild relatives* . *Nature Reviews Genetics* 4, 806-817)

## Bibliography

Bae JM, Ahn MY, Harn CH, Jeong M, Lim YP, Liu JR. Cloning and characterization of plastid ribosomal protein s16 gene from potato (*Solanum tuberosum* L. cv. Desiree). *Molecules and Cells* 8, 466–470. (1998)

Brisson LF, Zelitch I, Havir EA. Manipulation of catalase levels produces altered photosynthesis in transgenic tobacco plants. *Plant Physiology* 116, 259–269. (1998)

DANIDA, 2002: Assessment of potentials and constraints for development and use of plant biotechnology in relation to plant breeding and crop production in developing countries. Working paper. Ministry of Foreign Affairs, Denmark.

Datta K, Potrykus I, Datta SK: Efficient fertile plant regeneration from protoplasts of the Indica rice breeding line IR72 (*Oryza sativa*). *Plant Cell Rep* 11:229-233 (1992).

Dunwell JM. Novel food products from genetically modified crop plants: methods and future prospects. *International Journal of Food Science and Technology* 33, 205–213 (1998)

Feinberg AP, Vogelstein B: A technique for radio labeling DNA restriction fragments to high specific activities. *Anal Biochem* 13:6-13 (1983)

Kota M, Daniell H, Varma S, Garczynski SF, Gould F, Moar WJ. Overexpression of the *Bacillus thuringiensis* (BT) Cry2Aa2 protein in chloroplasts confers resistance to plants against susceptible and BT-resistant insects. *Proceedings of the National Academy of Sciences, USA* 96, 1840–1845 (1999)

Meijer EGM, Schilperoort RA, Rueb S, Os-Ruygrot PE, Hensgens LAM: Transgenic rice cell lines and plants: expression of transferred chimeric genes. *Plant Mol Biol* 16:807-820 (1991).

Murray HG, Thompson WF: Rapid isolation & high molecular weight plant DNA. *Nucleic Acids Res* 8:4321-4326 (1980).

Potrykus I: Gene transfer to plants: assessment of published approaches and results. *Annu Rev Plant Physiol Plant Mol Biol* 42:205-224 (1991).

Schell J: Transgenic plants as tools to study the molecular organization of plant genes. *Science* 237:1176-1183 (1987).

Toenniessen GH: Potentially useful genes for rice genetic engineering. In: Khush GS, Toenniessen GH (eds) *Rice Biotechnology*, pp. 253-280 C.A.B. International, UK/IRRI, Manila (1991).

Neal Stewart, Jr, Matthew D. Halfhill & Suzanne I. Warwick., ( 2003), Transgene introgression from genetically modified crops to their wild relatives .*Nature Reviews Genetics* 4, 806-817

## Module 6- Lecture 2

### BIOPHARMING - PLANTS AS BIOREACTORS

#### 6-2.1 Introduction

In previous chapter we studied about transgenic plants and a few techniques to develop them. In this chapter we will discuss about various transgenic plants developed so far for the therapeutic applications and how they can be used as bioreactors. The use of genetically engineered plants for the production of therapeutically important bio molecules is known as plant **Biopharming**. This is also called “molecular farming or pharming”. These plants are different from medicinal plants which are naturally available.

#### 6-2.2 Transgenic Crops for Therapeutic Application

- 1) **Transgenic rice:** There are few examples of transgenic rice developed for therapeutic purposes such as –
  - a)  **$\beta$ -carotene:** Golden Rice is capable of producing  $\beta$  –carotene (provitamin-A) in the endosperm (main grain) which is a pre cursor for vitamin A. The golden rice was developed by introducing one gene from daffodil and one gene from the bacterium *Erwinia uredovora* were into rice genome. The combination of these three genes help this plant to produce the enzymes necessary to convert geranyl-geranyl diphosphate (GGDP) to provitamin-A. The pathway of production of provitamin- A in golden rice is shown below:

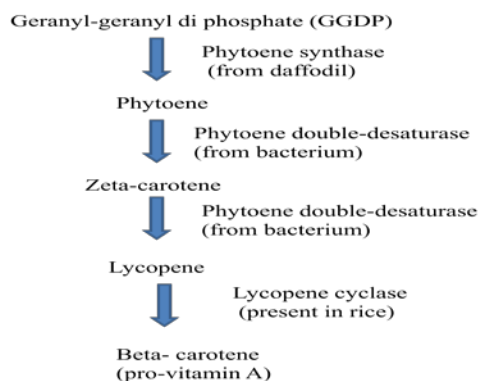


Figure 6-2.2.1 Pathway of production of provitamin A in golden rice

- b) Human Milk Proteins:** Rice has been as an expression system to express human milk proteins. Till today, the proteins lactoferrin, lysozyme and 1-antitrypsin expressed at very high levels, and large-scale field trials for several generations for the stability of these transgenic rice varieties are satisfactory.
- c) Transgenic rice containing high amount of Iron and Zinc:** A rice variety was developed to overcome the iron deficiency by introducing the soybean ferritin protein gene in rice genome. Iron deficiency is a major human nutritional disorder in the world. The ferritin protein takes up iron atoms and stores it in a non-toxic form. It releases iron when needed for metabolic functions. Finally it led to higher iron and zinc levels in transgenic rice grains.
- 2) **Transgenic potato:** A gene from Amaranth has been used in developing a variety with high protein content. Another variety was raised for producing antigens of cholera and diarrheal pathogens. Recently a genetically modified potato has been developed which carries the gene for the hepatitis B surface antigen.
- 3) **Transgenic maize:** Maize with genes from the 2G12 antibody (capable to neutralize infection from the HIV Type 1 Glycoprotein gp120) could produce antibodies against the transmission of HIV. Various transgenic maizes are also developed capable to produce AIDS antigens by expression of the SIVmac239 gp120 subunit (the major HIV envelope glycoprotein, gp120. Similarly some other maize varieties were also developed for higher content of lysine and tryptophan (essential amino acids) and for nutritive value equivalent to that of milk.
- 4) **Transgenic fruits and vegetables:** Various transgenic fruit and vegetables have been developed to produce subunit vaccines against Rabies, antigen for AIDS which can be used to develop antibodies. A tomato have been developed which can cure ulcer and stomach cancer by inhibiting *Helicobacter pylori*, which is a type of Gram-negative, microaerophilic bacterium located in the stomach.
- 5) **Transgenic tobacco:** Tobacco plant is the best choice for transgenics. Genes for the synthesis of human hemoglobin, human collagen and many others have been successfully introduced or are in the process of development. Recently a tobacco suspension culture which expressed recombinant human monoclonal antibody (MAb) against hepatitis B virus (HBV) surface antigen (HBsAg) has been developed.

### 6-2.3 Plant Made Industrial Products (PMIPs) on the Market:

➤ **Avidin**

Avidin is a tetrameric molecule and biotin which is a dimeric binding protein produced in the oviducts of birds, reptiles and amphibians. Functional avidin is found only in raw egg. It has affinity for biotin which is applicable in wide-ranging biochemical assays, including western blot, ELISA. Traditionally avidin isolated from chicken egg whites. It is now produced from transgenic corn.

➤ **GUS ( $\beta$ -glucuronidase)**

It is used as visual marker (i.e. reporter gene) in research labs. The gusA gene will show response by producing the enzyme  $\beta$ -glucuronidase (GUS) which turns Xgluc (a chromomeric substrate) into blue colored product after incubation for sometime. Traditionally it is isolated from bacterial sources (*E.coli*). It can be now obtained from transgenic corn.

➤ **Trypsin**

It has variety of applications, including biopharmaceutical processing. Traditionally trypsin (EC 3.4.21.4) was isolated from bovine pancreas. Now it is produced by transgenic corn and in fact is the first large scale transgenic plant product that gained commercial success.

### 6-2.4 Plantibodies

These are monoclonal antibodies produced in plants. Plants used include tobacco, corn, potatoes, soy, alfalfa, and rice. They are free from potential contamination of mammalian viruses.

- **Application in human:** Plantibodies are widely used to combat against several human diseases including cancer, dental caries, herpes simplex virus, respiratory syncytial virus.
- **Application in plants:** It involves protection from pathogens and pests. This is termed as immunomodulation. Plantibodies interfere with pathogen infectivity and block plant-pathogen infection.

Examples include –

ScFvs (Variable light and heavy chain fused with signal peptide) engineered against Artichoke Mottled Crinkle virus decreases infection and delay symptoms.

Full length anti-TMV antibodies decreased viral necrotic lesions.

ScFvs against Beet Necrotic Yellow Vein Virus gives a partial protection to the virus.

## 6-2.5 Production Costs of pharmaceutical from different expression systems

The production cost of pharmaceuticals from various systems is compared in table 6-2.5

Table 6-2.5 Production cost of pharmaceuticals from different systems

Production costs	Cost in \$ /gram
Yeast	50-100
Mammalian Cells	500-5000
Transgenic animals	20-50
Transgenic plants	10-20

Source: Data from Bio Pharm International (2003)

## 6-2.6 Comparison of Mammalian and Plant-produced Antibodies:

Heterologous protein production in plants has become incessantly efficient due to recent advances in plant biotechnology. These plant-produced monoclonal antibodies have been explored since they were expressed in tobacco several years ago. Both full length and partial antibodies produced in transgenic plants offer many intriguing possibilities to plant molecular biologists and plant breeders.

But polyclonal antibodies are typically produced in suitable mammal, such as a mouse, rabbit or goat. Larger mammals are often preferred due to their huge amount of serum. Antigen is injected into the mammal which induces B-lymphocytes to produce IgG immunoglobulin specific for antigen. This polyclonal IgG can be purified from the mammal's serum.

- The peptide sequence is identical.
- Correct cleavage of Ig-derived signal peptides takes place.
- Kinetics and affinity i.e. rate of association is identical.
- Stability of protein in seeds is more than 30 months.
- Plant system is more versatile in producing different types of antibodies like immunoglobulin A (IgA).



- Post-translational processing is different.

## 6-2.6 Advantages of Using Plants for Biopharming

- **Cost reduction**

- scalability
- low or no inputs
- low capital cost

- **Stability** – Storage of plant-derived products is possible and stable at room temperature or in a similar way of storing fruits and vegetables.

- **Safety**

- eukaryotic production system
- free of animal viruses (e.g. BSE)

## 6-2.7 Disadvantages of biopharming with plants:

- **Environment contamination**

- Gene flow to natural population.
- wildlife exposure

- **Food supply contamination** – mistaken or intentional mixing with human food.

- **Health safety concerns** - Variable, case-specific.

## 6-2.8 Stability and Storage

Proteins, carbohydrates, lipids and the other natural products in the seeds of crops can be stored well for decades. Therefore, drugs, vaccines and antibodies produced in crop plants are expected to be stable for long periods. Antibodies and vaccines are being well designed in such a way that it becomes edible in the raw condition to protect it from denaturation by cooking them.

## 6-2.9 Risks and Concerns

- **Environment contamination**
  - Gene flow via cross pollination (some agent like wind, insect etc.).
  - Target species near field sites, e.g. butterflies, bees, etc.
- **Food supply contamination** - Accidental, intentional, gene flow.
- **Health safety concerns**- It includes non-target organ responses, no side-effects, allergenicity.

## 6-2.10 Bio security Issues

Some biosecurity regulations of the transgenic non-biopharm crop will be applicable for biopharming crops. The therapeutic products regulations for the safety and efficacy will also be applicable on these transgenic plants. However, severe resistances of anti-tech groups raise against biopharm crops.

## 6-2.11 Proposed safeguards for biopharming operations:

- Physical differences - Easily detectable by addition of reporter genes e.g. “Purple” maize, GFP.
- Sterility which involves male sterile plants, somatic hybridization.
- Use of chloroplast expression system: It will help to increase yield, eliminate potential gene flow via pollen but has a limitation that it is technically difficult.
- Complete disclosure of DNA sequences.
- Legislate for administration.

## Bibliography

Akira Yano, Fumiko Maeda, Masataka Takekoshi. (2004) “Transgenic tobacco cells producing the human monoclonal antibody to hepatitis B virus surface antigen”; Journal of Medical Virology, 73(2):208–215.

Ashraf S, Singh PK, Yadav DK, Shahnawaz M, Mishra S, Sawant SV, Tuli R.(2005) High level expression of surface glycoprotein of rabies virus in tobacco leaves and its immunoprotective activity in mice. Journal of Biotechnology, 119:1-14.

Balen B, Krsnik-Rasol M, Lt. (2007) N-glycosylation of recombinant therapeutic glycoproteins in plant systems. Food Technology and Biotechnology, 45:1-10.

<http://www.in-pharmatechnologist.com/Regulatory-Safety/Plant-molecular-farming-set-to-boom>

Humphreys, J. and Chapple, C. (2000) "Molecular 'pharming' with plant P450s". Trends in Plant Science 5 (7): 271.

Keeler, K. H. (1989) "Can Genetically Engineered Crops Become Weeds?" BioTechnology. 7(11): 1134-1139.

Miller, H. I. (2003) "Will we reap what biopharming sows?" Nature Biotechnology 21 (5): 480–481.

Nguyen Thi Lang, Tran Anh Nguyet, Nguyen Van Phang and Bui Chi Buu, (2007) "Breeding For Low Phytic Acid Mutants In Rice (*Oryza Sativa L.*)"; Omonrice 15: 29-35.

Nilesh P. Teli & Michael P. Timko, (2004) "Recent developments in the use of transgenic plants for the production of human therapeutics and biopharmaceuticals" Plant Cell, Tissue and Organ Culture 79: 125–145.

Shihshieh Huang, Alessandra Frizzi, Cheryl A. Florida, Diane E. Kruger and Michael H. Luethy (2006) "High lysine and high tryptophan transgenic maize resulting from the reduction of both 19- and 22-kD  $\alpha$ -zeins"; Plant Molecular Biology.61:525–535.

Susan J. Karcher. (2002) "Blue Plants: Transgenic Plants with The Gus Reporter Gene"; Department of Biological Sciences, Purdue University, Chapter 3: 29-42.

## **MODULE 6- LECTURE 3**

### **TRANSGENIC SCIENCE FOR ANIMAL IMPROVEMENT**

#### **6-3.1Introduction:**

Transgenic science offers the methods for in vitro introduction of DNA fragment(s) into a cell and alters the genome in order to achieve a desired character. Animals whose genomes have been altered by transfer of a gene or introducing an exogenous gene are called transgenic animals. Transgenic animals serve as important model for clinical research as well as bioreactors in pharmaceutical manufacturing sector. The term “transgenic” was first coined by Gordon and Ruddle in 1981 and later adopted to describe broad range of experimental animal models having a specific targeted modification in their genome.

#### **6-3.2Types of transgenic animals:**

Transgenic animals can be broadly classified into two categories depending on the modification in the genome.

- **Knockout animals:**

Transgenic animals produced by replacing the endogenous functional gene with mutated non-functional gene are called knockout animals. In laboratory, knockout mice are generated for clinical studies.

- **Knockin animals:**

Transgenic animals having exogenous functional gene incorporated to the genome by targeted insertion are termed as knockin animals. Several animal transgenic models have been developed by this approach to use as disease models or in biofarming technology.

### **6-3.3 Methods for creating transgenic animals:**

With the advances of recombinant DNA technology, different methods have been developed for generating transgenic animals. The examples of commonly used methods applied in transgenic animal production are-

- a) Gene knockdown using RNA interference (RNAi),
- b) Retroviral mediated gene transfer
- c) Embryonic stem cell technology
- d) DNA microinjection
- e) Nuclear transfer
- f) Sperm as vector

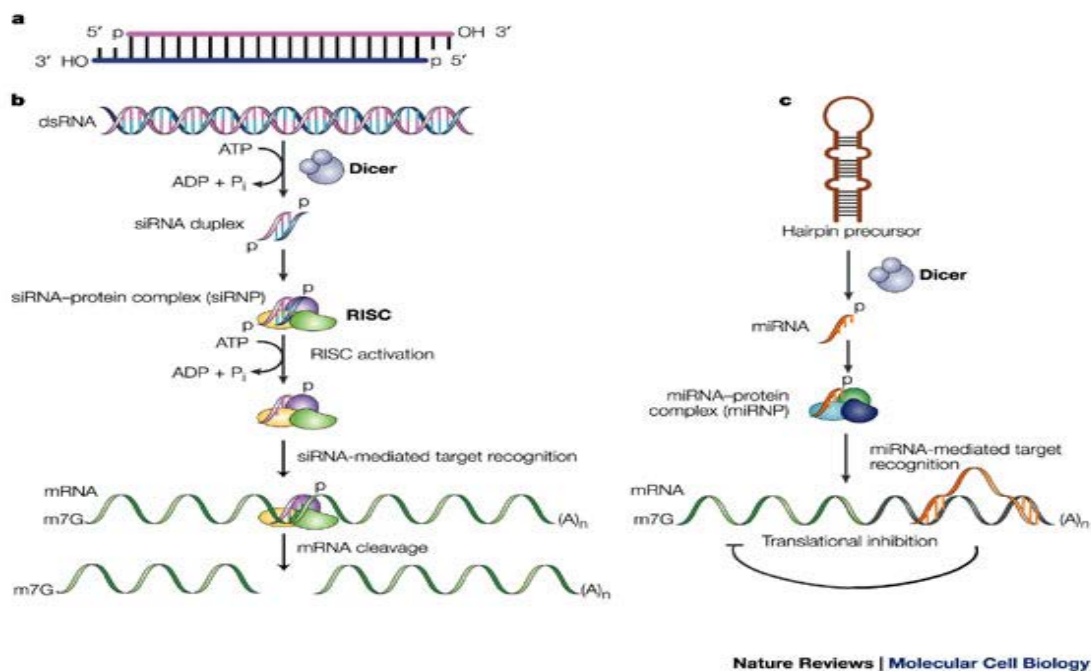
#### **6-3.3.1 Gene knockdown using RNA interference (RNAi):**

RNA interference (RNAi) induced in mammalian cells by introducing antisense RNA to generate double stranded RNA (dsRNA) structure (or hairpin) which are cleaved to generate small fragments of RNA (~22bp length). Double stranded RNA are identified and cleaved by Dicer enzyme to generate small interfering RNA (siRNA) or micro RNA (miRNA). miRNA is single stranded RNA molecule which inhibits translation of mRNA whereas siRNA are double stranded RNA which cleave RNA. A riboprotein complex is generated by the small RNA fragments called RISC (RNA induced silencing complex) which binds to homologous mRNA molecule and initiate the cleavage, thus silencing the gene expression. (Fig6-3.3.1)

In mammalian cells, RNAi effect can be introduced by generating small interfering RNA (siRNA) or small hairpin RNA (shRNA). Using this approach, several transgenic mice and rat models have been developed for potential medical and pharmaceutical applications.

RNAi technology has several advantages over other approaches. Once stably established RNAi can be synthesized directly inside the mammalian cells thus avoiding any other

cloning steps. RNAi approach can be designed to regulate gene expression either transcription or translation stage.



**Fig 6-3.3.1: Gene silencing by RNAi: a. Double stranded RNA construct formed by introducing antisense RNA into the system; b. Mechanism of generation of siRNA from dsRNA by Dicer in presence of ATP; c. Generation of miRNA from hairpin precursor RNA (shRNA) by Dicer.**(Source: [http://www.nature.com/nrm/journal/v4/n6/fig\\_tab/nrm1129\\_F1.html](http://www.nature.com/nrm/journal/v4/n6/fig_tab/nrm1129_F1.html), reprinted with permission.)

### 6-3.3.2 Retroviral mediated gene transfer:

Retrovirus is a virus of family “Retroviridae” that has RNA as its genetic material and replicate inside the host cell using reverse transcription machinery (to convert RNA to DNA). Retroviruses are commonly used as vectors to transfer genetic material into the host cell because of its high efficiency to transfer RNA to the host. Recombinant retrovirus RNA genome is then copied by reverse transcriptase (encoded by retrovirus) to yield a DNA copy, which then becomes integrated into the host genome at random sites. The offspring results from such transformation are chimeric, an organism consisting of tissues of diverse genetic constitution. To transmit transgene to next generation, any gene transfer technique including retroviral mediated should target germline cells. Generally, early 4-16 celled embryos are used to get maximum of retrovirus infection resulting in chimeric animal. Although homozygosity is attained just after several generations of

selfing, these chimeric animals are inbred for 10-20 generations to obtain homozygous transgenic animals.

Although retroviral vectors have high efficiency of transgene integration into host cell, there are certain limitations of this approach.

**Limitations:**

- Low copy number integration.
- Additional steps needed for construction of recombinant retrovirus.
- Limited size of DNA insert (~<15kb).
- Mosaicism (same individual having two different types of genotypes) of the recovered animals.
- Possible interference of the proviral LTR sequences with the expression of transgene.

### 6-3.3.3 Embryonic stem cell technology:

Embryonic stem (ES) cells are pluripotent stem cells isolated from inner cell mass (ICM) of blastocyst stage of embryo. Use of ES cell mediated transgenic animal production is quite effective specially to generate targeted gene modification at precise location.

The procedure to generate transgenic animals using modified ES cell technology is described in brief below:

- i. Isolation of pluripotent ES cells from ICM of blastocyst and culturing of the cells *in vitro*.
- ii. Transgene introduction to ES cells using methods like DNA micro-injection, electroporation, precipitation reaction, transfection etc. DNA transfection to ES cells can be carried out using different vectors like liposome, retroviral vectors etc.
- iii. Selection of transformed ES cells for either knockin or knockout constructs.

- iv. Transformed cells are injected into blastocyst followed by implantation in a surrogate female.
- v. After birth, the offsprings are screened for chimerism. Inbreeding of the genetic chimeras is performed to obtain homozygous transgenic animal carrying both the mutated alleles for that character.

**Limitation:**

- Difficulty in *in vitro* culturing and maintenance of pluripotent ES cells.
- Germ line transfection with low frequency.

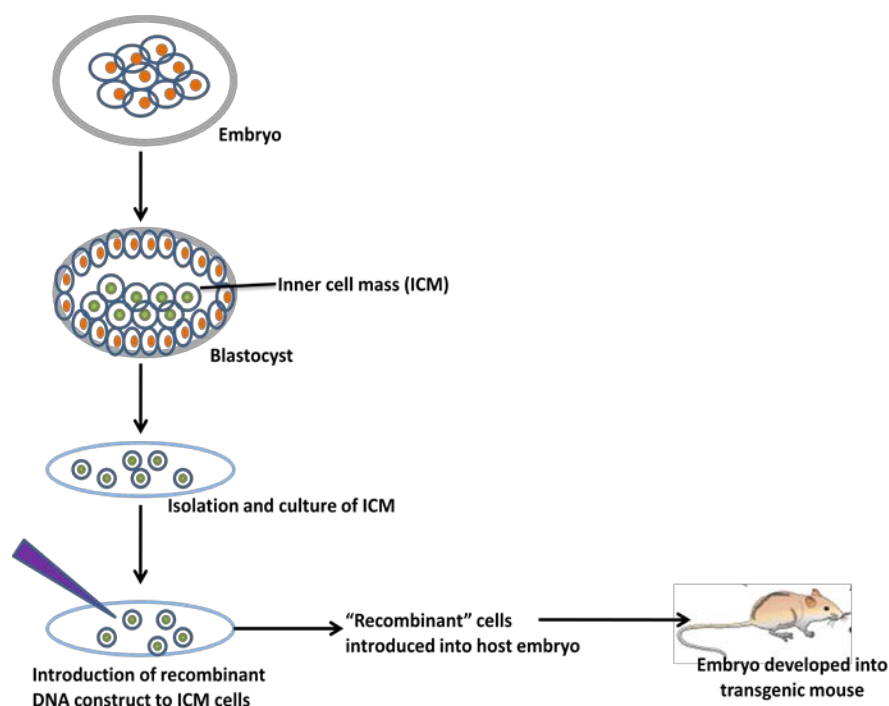


Fig 6-3.3.3: Embryonic stem cell technology

#### 6-3.3.4 DNA Microinjection:

DNA microinjection technique was first described in 1966 by T. P. Lin, which paved the way of generating transgenic animal by direct introduction of foreign DNA into the host cell. Linearized DNA construct is injected into the cell which finally integrates to the host genome at random sites by homologous recombination. DNA microinjection can be targeted to two sites: **Pronuclear injection** and **cytoplasmic injection**. Following steps are involved in production of transgenic animal by microinjection:



**1. Transgene DNA Construct:** cDNA of a gene of interest along with the regulatory elements (promoter, terminator etc) are used to design the DNA construct. Promoter and 3' UTR are necessary for proper gene regulation. Some evidence shows that intron splicing plays role in gene expression in mammalian system. Hence introns are also included in the construct sometimes. Construct should be linearized before injection.

**2. Embryo Collection:** Donor parental strain for the production of embryo are selected considering factors like response to super ovulation (ability to produce mature ova at large number), frequencies of embryo survival following microinjection, size of pronuclei and incidence of specific pathologies inherent to strain.

Successful super ovulation depends on the strain, age and weight of the animal. Breeding should be monogamous. After post-mating, embryos are collected at single cell stage and proceed for microinjection.

**3. DNA Microinjection:** The purified double stranded linear DNA construct containing the transgene DNA sequence is introduced to host cell by microinjection. The foreign DNA must be integrated (although random) into the host genome prior to cell division. To facilitate this transgene DNA is introduced into zygote at the earliest possible stage (pronuclear period) immediately after fertilization. Usually male pronucleus is preferred because it is larger and easier to inject. The host chromosome at the site of integration generally undergoes duplication, deletion or rearrangements due to transgene incorporation. This may lead to insertional mutagenesis and thus producing detectable phenotypic trait.

**Note:** Critical points for successful DNA microinjection technique-

- Careful collection of relatively large group of accurately single cell embryo.
- Embryo transfer to suitable recipient female (standardized in each case).
- Construction and preparation of transgene DNA fragment to be injected.

**Advantage:**

- No size limitation of insert DNA.
- Applicable to wide range of species.
- Less time consuming and short generation time.

**Limitation:**

- Low success rate
- Random integration of transgene constructs to host chromosome, resulting in undesirable effect or phenotype.
- Potential undesired insertional mutagenesis

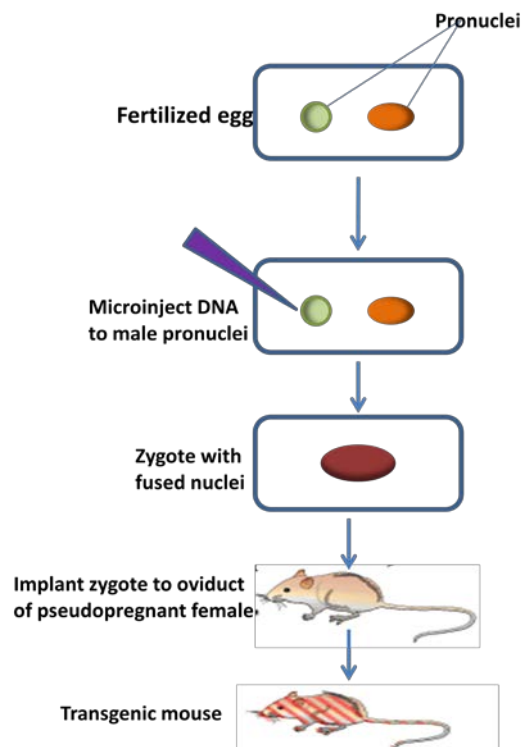


Fig 6-3.3.4: Transgenic mice by DNA microinjection

**6-3.3.5 Nuclear Transfer:**

Nuclear transfer involves transfer of nuclei from a donor cell (somatic cell or stem cell) to an enucleated oocyte, thereby reprogramming development. This process is also referred as somatic cell nuclear transfer (SCNT). In early 1980, first mammalian nuclear transfer experiments were conducted in mice. In 1997, a team lead by Ian Wilmut created

Dolly the sheep became the first vertebrate clone to generate by somatic cell nuclear transfer. The technique however was not full proof as Dolly was reported to have progressive lung disease and premature arthritis and died prematurely at the age of 6 in 2003.

Following steps are associated with generation of transgenic animal by SCNT procedure:

1. Mature oocyte is arrested at metaphase-II phase and enucleated.
2. Somatic donor cells are cultured in vitro and somatic cell derived nucleus is transferred to enucleated oocytes.
3. The resulting nuclear transferred oocytes are activated and embryogenesis is activated.
4. The embryo is generally allowed to develop in vitro till blastocyst stage.
5. The blastocyst embryo is then implanted into the uterus of surrogate female where the further development of the embryo takes place.

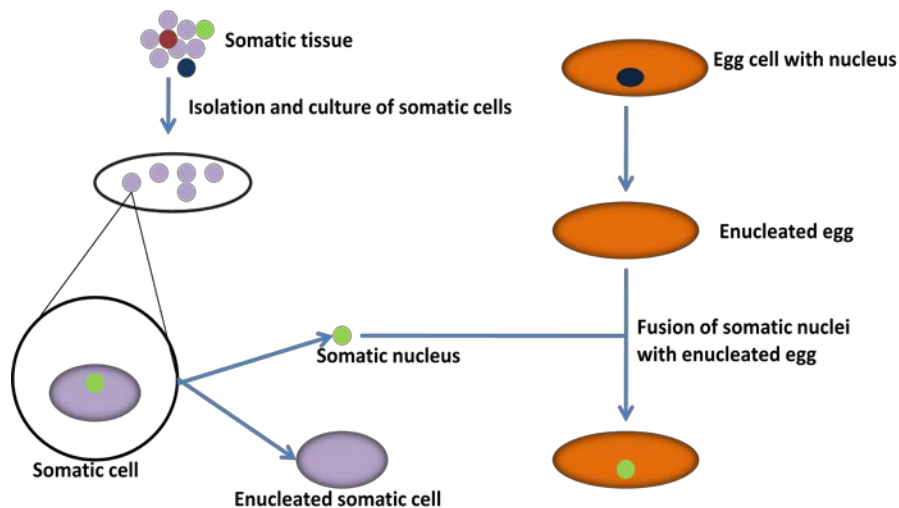


Fig6-3.3.5 Somatic cell nuclear transfer

### 6-3.3.5 Sperm Mediated Gene Transfer (SMGT):

Sperm mediated gene transfer was first reported in 1989 which is based on ability of sperm cells to internalize exogenous DNA and transfer it to the egg cells. This method utilizes a “natural vector” (sperms) to deliver foreign DNA. The exogenous DNA binds to sperm’s head at sub-acrosomal region and DNA molecules are taken up by sperm cells. After DNA-sperm interaction and internalization, the exogenous DNA gets integrated to the genome.

#### **Criteria for Successful SMGT Technique:**

- Quality of semen sample
- DNA uptake rate which depends on viability and motility of sperm cells.

#### **Limitation:**

- Sperm cells take up any DNA from the environment which reduces the specificity of the vector
- High amount of foreign DNA inhibits sperm movement.
- Large amount of DNA may get fragmented during the activation of sperm nucleus.

### **6-3.4 Characterization of Transgenic Animals:**

After successful introduction of transgene into the host animal, the resulting transgenic animal should be kept under observation to study the stability and expression pattern of transgene inside host system. Continued expression of the product depends on two aspects of stability of the transgene:

- The structural stability or integrity of the transgene.
- The stability of transgene expression.

### **A. Genetic Stability:**

In presence of multiple integration sites in the host genome, there may be occurrence of rearrangement or deletion of integrating transgene as well as host genome. To overcome this, stability of the transgene should be regularly monitored by tests like Southern blots, sequencing etc

### **B. Stability of expression:**

The stability of expression of a transgene product can vary depending on interaction of the genetic background of the host animals or imprinting effects due to maternal or paternal inheritance. Stability of expression of the transgene product should be monitored over the productive life of the transgenic animal. Methods may include Northern blots, RT-PCR, DNase protection assay or other appropriate techniques.

- Detailed characterization of the original gene intended for introduction into the animals should be provided.
- The natural protein and its functions should be described along with a description of its pattern of expression.
- Both the original vector sequence and the transgene construct should be extensively characterized by restriction maps and nucleotide sequences.

### **C. Characterization of the transgenic founder (F0) animal:**

- The medical history of the animals which donate gametes or embryonic stem cells (donor) and foster or recipient animals should be well known.
- Animals should undergo detailed veterinary evaluations of health, including breed-related and specific tests for species related disease problems.

### 6-3.5 Importance of Transgenic Animals:

- 1. Targeted Production of Pharmaceutical Proteins:** Transgenic animals are used for the production of valuable human enzymes, hormones, antibodies and growth factors. Transgenic animals (eg. larger mammals) as **bioreactor** (“Pharmaceutical pharming”) is cost effective compared to cell culture methods as animals can supplement their body fluid with fresh nutrients, remove waste products, internally regulate the temperature and pH and resist pathogens on their own. The transgene products can be collected by targeting its expression in the secretory cells of the liver, lactating mammary gland and kidney. e.g. - Sheep secreting human Factor IX in milk used for the treatment of hemophilia B.
- 2. Models of Human Disease Processes:** To study different human disease patterns and their mechanisms, different lower animals having evolutionary resemblance to human are used. Due to stringent regulatory and ethical issues, direct experiment on human is not possible. The phenotypic and regulatory parameters of a particular human disease may be evaluated in an animal model with relatively short generation time. By expressing human proteins in a transgenic animal, the regulation and activity can be studied. Also by creating a condition similar to human diseased condition (eg. diabetic mice), one can study various parameters and their interaction. Developing diverse transgenic phenotypes permits the innovative testing of diagnostics and therapeutic agents with a reduced population of experimental animals. Utilizing novel cell lines from transgenic organs can lead to huge reduction in the number of research animals required for the evaluation. It is possible to create a transgenic genome in which more than one transgene may interact or in which a transgene may interact with endogenous normal or mutated gene. e.g. - transgenic rodent model are developed for numerous human disease like cardiovascular disease, cancer, etc.
- 3. Analysis of a Developmental Pathway:** Many transgenic animals are non-viable, either prenatally or perinatally. This can give insights into genes essential for during development. Correlation of time of embryonic lethality to known animal embryonic

gene expression patterns provides knowledge of insight into the pathways and interaction of the mutated gene. Information on gene expression can easily be obtained by using a reporter gene, which combines the regulatory regions of the gene under study with an easily detectable marker protein. For e.g. – Using Cre-lox system both temporal and tissue specific control of expression of knockout gene construct can be studied.

- 4. Modification of Animal Anatomy and Physiology:** Transgenic technology can be used for genetic improvement of an animal. A foreign gene can be introduced by selective deletion of endogenous gene or specific region. A direct approach is to add a gene for a growth factor or hormone to the genome instead of altering the complex multigenic physiology of the mammal. For developing animals with novel disease resistance, higher yield of meat or milk, decreased body fat, etc.
- 5. Production of Organs for Xenotransplantation:** Transplantation of living cells, tissues or organs from one donor organism to another recipient organism is known as xenotransplantation. Such cells, tissues or organs are called xenografts or xenotransplants. The shortage of healthy donor organ can be mitigated by developing animals whose organs would be suitable for xenotransplantation. Transgenic animals have aided the understanding of molecular basis of graft rejection, which includes foreign antigens, coagulation defects and interspecies ligand-receptor interactions. e.g. - Transgenic pigs have been developed that do not express the surface glycoproteins highly immunogenic to humans.

### Bibliography:

Brackett B.G., Baranska W., Sawicki W., Koprowski H. (1971); Uptake of Heterologous Genome by Mammalian Spermatozoa and Its Transfer to Ova through Fertilization; *Proc. Natl. Acad. Sci.* 68:353-357.

Bradley A., Evans M.J., Kaufman M.H., Robertson E.J. (1984); Formation of germ line chimaeras from embryo-derived teratocarcinoma cell lines; *Nature* 309: 255-256.

Derek M. Dykxhoorn, Carl D. Novina and Phillip A. Sharp (2003); Killing the messenger: short RNAs that silence gene expression; *Nat Rev Mol. Cell Bio.* 4, 457-467.

Gossler A., Doetschman T., Korn R., Serfling E., Kemler R.(1986); Transgenesis by means of blastocyst-derived embryonic stem cell line; *Proc. Natl. Acad. Sci.*83:9065-9069.

Jaenisch R., Mintz B. (1974); Simian virus 40 DNA sequences in DNA of healthy adult mice derived from preimplantation blastocysts injected with viral DNA; *Proc. Natl. Acad. Sci.*71: 1250–1254.

Lavitrano M., Camaioni A., Fazio V.M., Dolci S., Farace M.G., Spadafora C. (1989); Sperm cells as vectors for introducing foreign DNA into eggs: genetic transformation of mice; *Cell*57:717-23.

M. Lavitrano et al (2006); Sperm mediated gene transfer; *Reproduction, fertility and development* 18:19-25.

Mephram T. Ben et al; The Use of Transgenic Animals in the European Union: the Report and Recommendations of ECVAM Workshop.

Richa J., Lo C.W. (1989); Introduction of human DNA into mouse eggs by injection of dissected chromosome fragments; *Science*245: 175-177.



## **MODULE 6-LECTURE 4**

### **BIOPHARMING-ANIMALS AS BIOREACTOR FOR RECOMBINANT PROTEIN**

#### **6-4.1 Introduction:**

As discussed earlier, transgenic animals are genetically modified by stable incorporation of foreign DNA into its genome by artificial gene transfer technology in order to introduce, silence or delete specific phenotypic characteristic. Transgenic animals can be used to express recombinant proteins in milk, urine, blood, sperm, or eggs, or even to grow organs for xenotransplantation. “Biopharming” in animals encompass the production of biologically active recombinant proteins using transgenic animals as bioreactor.

Although production of recombinant human therapeutic proteins in microbial bioreactors is extremely cost-efficient, it suffers from many limitations like:

1. Prokaryotic cells lack post translational modification machinery (glycosylation, phosphorylation etc) essential for proper folding and functioning of complex eukaryotic proteins.
2. The lack of post-translational modifications in prokaryotic system can be overcome by replacing them with large-scale animal cell cultures. But the long generation time and the requirement for rich culture media make animal cell bioreactors more expensive.
3. In both the above cases, the upstream production operations are relatively inexpensive but the downstream processing of the final product is sometimes extremely complex and costly.

These drawbacks can be overcome by using transgenic animals for production of human or veterinary pharmaceuticals. The use of targeted gene transfer for the expression of the transgene in the mammary gland of large farm animals like cow, pigs, goats and sheep offers advantages like easy isolation of the product secreted into the milk and low cost-large scale production.

Transgenic have been produced in a variety of animal species e.g., mice, rabbits, swine, sheep, goat, cattle, poultry, fish, amphibians, insects and nematodes to produce therapeutic proteins against osteoporosis, cystic fibrosis, hemophilia, malaria, arthritis, HIV etc. Transgenic animals have also been developed to produce monoclonal antibodies (therapeutic antibodies against disease proteins) which are used in vaccine development. Other than being used as bioreactor, transgenic animals are also used as disease models and for improving livestock.

### 6-4.2 Advantages of Using Transgenic Animal in Biopharming:

- Protein function and activity depends on the three dimensional structure achieved by **post-translational modification**. The post-translational modification machinery is absent in bacteria (like *E. coli*) and plants have significantly different machinery than animals. To produce complex human therapeutic proteins, nothing can be better than animal livestock.
- Animal pharming does not require **constant monitoring** and **sampling**.
- Animal pharming is **cost effective** as it does not require substantial plant machinery which has to be purchased and maintained otherwise.
- **Purification of proteins** from an animal's secretory or excretory gland products (milk, body fluid, urine etc) is simpler.
- **Scale-up** of production of a recombinant protein is easier with animal livestock.

### 6-4.3 Purification and Characterization of Transgenic Product:

#### 6-4.3.1 Product recovery from the transgenic animals:

Recovery can be done by milking, exsanguinations (process of blood collection), surgical excision of tissues, or other methods as applicable. The recovery procedure should be designed to maximize the safety, sterility, potency, and purity of the product. The collection facilities should be maintained at the best possible hygiene condition to reduce any contamination.

### 6-4.3.2 Quality of a production batch:

Monitoring the quality of the product and the manufacturing processes require the purity of the product to be determined at multiple points during the production process. The tests found to be most sensitive to changes in product structure and/or potency should be validated and then incorporated as part of the batch release protocol. Conditions and criteria by which a production batch is defined (e.g., animals used, batch size, collection storage times, the time before sterile filtering and acceptance criteria for source material and product pooling) should be provided.

### 6-4.3.3 Endogenous and Adventitious Agents Detection:

#### 1. Choice of host animal:

Many animal species have been proposed to develop as hosts for production of therapeutic products. The lack of experience with many of these hosts raises potential safety concerns about adventitious agents are considered on a case-by-case basis by the FDA. **Adventitious agents** are defined as micro-organisms which are unintentionally introduced to the manufacturing process of a biological product. There is a possibility that adventitious agents or chemical contaminants could enter into the host animals and be concentrated in the product by the purification procedure without signaling their presence. Zoonotic diseases (diseases caused by infectious agents that can be transmitted from vertebrate animals to humans and vice-versa) should be taken care of extremely. Therefore, health monitoring of the livestock is necessary, although not sufficient alone to guarantee absence of these contaminants. The rigor of infection control in the animal host and validation of elimination of adventitious agents from the product will depend on several factors like:

1. The intended use of the product;
2. The tissue from which the product is derived;
3. The method by which product is collected;
4. The purification process;

5. Animal husbandry practices used during production of the founder and production animals.

## **2. Source tissue:**

Currently therapeutic products are produced in transgenic animals by utilizing systems in which the product will be isolated from bodily fluids (milk, blood, or urine) where significant batch-to-batch variation in both microbiological and virological loads may arise in unpurified products. The extent of such variability should be documented for each product. The downstream manufacturing procedures should have full capability for providing a safe and consistent product.

## **3. Pathogen testing and elimination:**

Transgenic animal source materials have the potential to contain a variety of human pathogens (e.g., viruses, bacteria, mycoplasma, and transmissible spongiform encephalopathy agents). Analytical methods and purification processes should be standardized to ensure the safety of the product. After consulting with veterinarians and other experts in the field, a specific list of pathogenic agents and accepted testing methods should be prepared and prioritized with regard to the risk of human infection. In cases where recombinant DNA was introduced by viral vector administration, or similar methods, the possibility of viral load in the product should be monitored.

### **6-4.3.4 Analysis of product identity and purity:**

Before initiating studies in humans, the product from the transgenic source material should be characterized with respect to activity, safety, purity, and potency. Product similarities (e.g., potency, pharmacokinetics) and differences (e.g., glycosylation, antigenicity) with well characterized natural or recombinant molecules should also be performed.

#### **Biochemical identity, purity and potency:**

Physicochemical, immunological and biological characterizations should be carried out on the transgene product to assess its identity, purity and activity which include chemical structure, amino acid composition, disulfide linkages etc. The anatomical site of

production in the transgenic animal may differ from the natural site of production and result in biochemical modifications that affect the activity of biological product.

**Product heterogeneity:**

Variations in final product activity should be examined with respect to variables associated with maintaining the transgenic herd and generating source material. For example, the identification of any variation in product structure, purity or activity when the source material is obtained: a) from different animals, b) during different seasons, c) from animals facing other sources of stress, d) from source material that has been stored for extended periods of time etc should be ascertained.

Characterization of the transgenic production process should initially focus on:

- 1) The presence of known and potential human pathogens, (both adventitious and endogenous agents) in the source material.
- 2) The amount of immunogenic and toxic materials in the final product,
- 3) The batch-to-batch consistency of the products physicochemical properties.

**6-4.3.5 Batch release testing of the final product:**

The purified product prior to final formulation should be characterized in a manner similar to other recombinant products. Contaminant levels in the final product that are acceptable for product release will be determined by preclinical studies, the dose and route of product administration, frequency and duration of product administration and the proposed patient population for the clinical study.

**6-4.3.6 Preclinical safety evaluation**

If there is similarity between the product and the naturally occurring molecules, it may still require limited toxicology studies. *In vitro* studies in animal models may be used to examine product efficacy and safety.

## 6-4.4 Application of Transgenic Animal Biopharming

### 6-4.4.1 Biomedical application:

To overcome the limitations of conventional recombinant production system of human therapeutic proteins, transgenic animals have been developed as bioreactor.

#### 6-4.4.1.1 Therapeutic protein production:

Several human therapeutic proteins were produced using transgenic livestock as bioreactor. Low cost, higher level of production and easy downstream processing make transgenic animal a better alternative to conventional production system. Table 6-4.4.1.1 shows examples of some therapeutic proteins produced using transgenic animals.

**Table 6-4.4.1.1: Human therapeutic proteins in transgenic animals**

<b>Animal</b>	<b>Drug/protein</b>	<b>Use</b>
sheep	alpha1 anti trypsin	deficiency leads to emphysema
sheep	CFTR (Cystic fibrosis transmembrane conductance regulator)	treatment of cystic fibrosis
sheep	tissue plasminogen activator	treatment of thrombosis
sheep	factor VIII & IX (blood clotting protein)	treatment of hemophilia
sheep	fibrinogen	treatment of wound healing
pig	tissue plasminogen activator	treatment of thrombosis
pig	factor VIII, IX	treatment of hemophilia
goat	human protein C	treatment of thrombosis
goat	antithrombin 3	treatment of thrombosis

goat	glutamic acid decarboxylase	treatment of type 1 diabetes
------	-----------------------------	------------------------------

#### 6-4.4.1.2 Antibody production using transgenic animals:

Mammary glands of several transgenic animals are targeted for specific production of monoclonal antibodies specific for various diseases. Transgenic goats and cattle are developed for such application.

#### 6-4.4.1.3 Blood replacement:

Functional human blood proteins including haemoglobin were produced using transgenic swine. Further improvement of this strategy is required to achieve desirable level of expression without any health risk to the animal and the recipient (human).

#### 6-4.4.1.4 Xenotransplantation of organs:

Organ transplantation has tremendous importance in medical science which is a critical life saving option in several cases. The shortage of human organs for transplantation leads to the development of xenograft techniques. There are certain pre-requisites for successful xenotransplantation:

- Overcome immunological rejection (called as “graft rejection”) of donor organs in the recipient body.
- Prevent transmission of pathogens from donor to recipient

Transgenic animals (pigs) have been developed to achieve reduced immunological rejection and better compatibility.

#### 6-4.4.2 Agricultural application:

Transgenic animal has been developed to provide various biopharming applications in agriculture. Few examples of the applications are discussed below-

##### 6-4.4.2.1 Improvement of economically important traits in livestock:

- Transgenic pig bearing human metallothionein promote/porcine growth hormone gene showed better growth rate, feed conversion, body fat/muscle ratio etc, thus improving the productivity.
- Transgenic pigs having human insulin-like growth factor-I had more lean muscle and about ~30% less fat; thus improving the commercial significance of these transgenics.
- Transgenic poultry animals were developed to increase the healthiness and nutrient quotient of meat and other products (less saturated fat, lactose level etc.).

##### 6-4.4.2.2 Lactation:

Dairy farming is an attractive field in agriculture which has huge commercial potential. Transgenic mice models have been developed to demonstrate the feasibility of such modifications in larger animals but further research is required to develop the concept. Production of milk having modified properties like lactose free milk, modified lipid content etc would have a niche market. Transgenic goat has been developed to secrete spider silk in goat milk.

##### 6-4.4.2.3 Wool production:

Transgenic sheep has been developed to increase wool production (~6.2% increase) by introducing “keratin-IGF-I” construct. Such modification showed no adverse effect on the growth and development of the transgenic animals.



#### 6-4.4.2.4Eco-friendly farm animals:

The high phosphorous content in manure of farm animals causes environmental pollution. Transgenic pigs have been developed bearing a bacterial “*phytase*” gene which allow digestion of plant phytate. Transgenic pigs showed ~75% lesser fecal phosphate output compared to the wildtype pigs.

### 6-4.5 ATryn® -First Recombinant Human Antithrombin:

The first product derived from a transgenic animal is a recombinant antithrombin called ATryn® was developed by GTC Biotherapeutics and approved by US Food and Drug Administration (FDA) in 2009. It is used for the prophylactic treatment of deep vein thrombosis in patients with hereditary antithrombin deficiencies. The promoter region from goat beta-casein gene was linked to human antithrombin (*hAT*) cDNA and the recombinant construct containing the transgene was introduced into the goat embryos, which were then implanted into surrogate mothers. The resulting offsprings carrying this transgene produced the recombinant human antithrombin (rhAT) in their milk. The rhAT protein is isolated from the milk by conventional purification method using tangential flow filtration, anion exchange chromatography, heparin affinity chromatography, nanofiltration and hydrophobic interaction chromatography with a yield of greater than 50%.

#### Important Safety Information:

- ATryn® is not advisable to use in patients with known hypersensitivity to goat milk proteins. Allergic-type hypersensitivity reactions, including anaphylaxis may occur in some patients. Use of drug must be discontinued at once if these reactions are reported.
- The alteration of anticoagulant effect of drugs that use antithrombin as anticoagulating agent may be observed when ATryn® is added or withdrawn. To avoid insufficient or excessive anticoagulation, suitable coagulation tests for the anticoagulant should be performed.

#### **6-4.6 Limitation of transgenic animals in pharming:**

- Both the time and cost to develop a transgenic model with high production level of therapeutic protein is high.
- Highly skilled persons are required for handling transgenic animals.
- Environmental and social impact of using transgenic animals as biopharming tools is complex and uncertain.
- Regulatory processes associated with the commercialization of a given product require description of detailed procedures and for the enforcement of intellectual property rights.
- Other concerns surrounding these technologies include the environmental and ethical aspects of transgenesis. Environmental issues associated with genetically modified organisms that have caused the most public outcry, including food and ecosystem contamination or threats to biodiversity, are generally more problematic in transgenic plants than animals.

#### **6-4.7 Safety Aspects of Biopharming:**

FDA has developed several regulations and guidelines for safe production and commercialization of recombinant products using transgenic animals. A crucial criterion in animal derived products is prevention of transmission of pathogenic factors from animals to human. All transgenic application of livestock requires compliance to the standards of genetic security and reliability of method applied in genetic modification.

## Bibliography

Hammer R.E., Pursel V.G., Rexroad C.E., Jr.Wall R.J., Bolt D.J., Ebert K.M., Palmiter R.D., Brinster R.L.(1985); Production of transgenic rabbits, sheep and pigs by microinjection; *Nature* 315:680.

Houdebine LM. (2000); Transgenic animal bioreactors. *Transgenic Res.*; 9(4-5):305-20. Review.

Hoyer, L. W., Drohan, W. N. And Lubon, H. (1994); Production of human therapeutic proteins in transgenic animals; *Vox Sanguinis*, 67: 217–220.

Jänne J, Hyttinen JM, Peura T, Tolvanen M, Alhonen L, Halmekytö M. (1992); Transgenic animals as bioproducers of therapeutic proteins. *Ann Med.* 24(4):273-80.

Palmiter R.D., Brinster R.L., Hammer R.E., Trumbauer M.E., Rosenfeld M.G., Birnberg N.C., Evans R.M.(1982); Dramatic growth of mice that develop from eggs microinjected with metallothionein-growth hormone fusion genes; *Nature*300: 611.

Schnieke A.E., Kind A.J., Ritchie W.A., Mycock K., Scott A.R., Ritchie M., Wilmut I., Colman A., Campbell K.H.(1997); Human factor IX transgenic sheep produced by transfer of nuclei from transfected fetal fibroblast; *Science*278:2130.

## **MODULE 6- LECTURE 5**

### **GENE MAPPING IN PLANTS AND ANIMALS**

#### **6-5.1 Introduction**

Gene mapping is like preparing a road map in which markers can be considered as the various milestones of the road between which lie the genes. The development of genetic map of an organism gives a picture of arrangement of genes in their chromosomes. The maps are composed of markers that may be genes controlling visible phenotypic traits (classical markers) or molecular markers whose phenotype is revealed by using modern molecular biology techniques (e.g. DNA markers). The goal of gene mapping is to study the regulation and expression of genes.

#### **6-5.2 Importance of Gene Mapping**

- Mapping helps in finding the inheritance of many rare genetic disorders such as cystic fibrosis, haemophilia etc.
- It helps in understanding the expression and regulation of a commercially important trait.
- The information generated by a map can be further utilized for marker assisted selection, generation of high quality breed of plant and animals.

#### **6-5.3 Types of Gene Mapping**

Gene mapping methods are divided into a) Genetic Mapping and b) Physical Mapping. In this lecture we will discuss in length about genetic mapping and briefly about physical mapping techniques.

##### **6-5.3(a) Genetic Mapping**

It uses genetic technique (such as pedigree analysis or breeding experiments) to design a map which can show the location of genes and the patterns of sequences in a genome. The basis of genetic mapping is Linkage analysis:

- The inheritance of the chromosomes is always as intact units where the linked alleles of paired genes will be inherited together because they are on the same chromosome.
- If a pair of gene is inherited independently then it is expected that either the genes are on different chromosomes or not linked although in same chromosome which is why they are sometime inherited together and sometime separately due to crossing over.

Genes (markers or loci) segregate via chromosome recombination during meiosis (i.e. sexual reproduction), thus allowing their analysis in the progeny. Chromosomes assort randomly into its gametes under meiotic cell division and these alleles segregate independently (according to Mendel's law of independent assortment). When two genes are close together on the same chromosome and they do not assort independently, they are said to be linked. At the beginning of meiosis, a homologous chromosome pair may intertwine and exchange sections of chromosome. Such process or set of processes is called **recombination** (also called **cross-over or strand exchange**). It results in two types of gametes:

- i) No crossing-over → **parental gametes**
- ii) Crossing over → **recombinant gametes**

Genes that are located on different chromosomes assort independently (unlinked) and have a recombination frequency of 50%. However, linked genes have recombination frequency of less than 50%.

### **Importance of Genetic of Mapping**

- To develop appropriate mapping population and decide the sample size.
- Selecting the type of molecular marker(s) to genotype the mapping population
- Screening of parents for marker polymorphism followed by genotyping the mapping population (parents and all progenies).
- To perform the linkage analyses (i.e. calculating pair wise recombination frequencies between the markers, establishing linkage groups, estimating map distances, and determining map order) via statistical programs.

### 6-5.3(a).1 Molecular Marker's for Gene Mapping

There are many molecular markers which can be used as detection systems for genetic variations by using genomic DNA for genetic analysis. Among these techniques, Restriction fragment length polymorphisms (RFLP), Random amplification of polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), Simple Sequence Repeat (SSR), and Single Nucleotide Polymorphisms (SNP) are most commonly used in genetic linkage analysis.

#### 6-5.3(a).1.1 Restriction fragment length polymorphisms (RFLP)

- RFLPs based molecular markers are amongst the first generation molecular markers used for many animal and plant genome mapping projects.
- This technique is based on digesting genomic DNA with a certain restriction enzymes and followed by southern blotting.
- Due to mutation there is a chance of gain or loss of restriction sites.
- As a result when homologous chromosomes are digested with restriction enzymes, electrophoresis and probed fragment length polymorphism is obtained (Figure 6-5.3(a).1.1).
- RFLPs are so sensitive that it can detect single- nucleotide mutations, deletions or insertions of DNA fragments of very small number of base pairs (1-100bp).
- Rearrangements of DNA on large chromosomal fragments can also be detected by RFLPs.
- RFLPs are well known for its robustness and accuracy in genotyping.

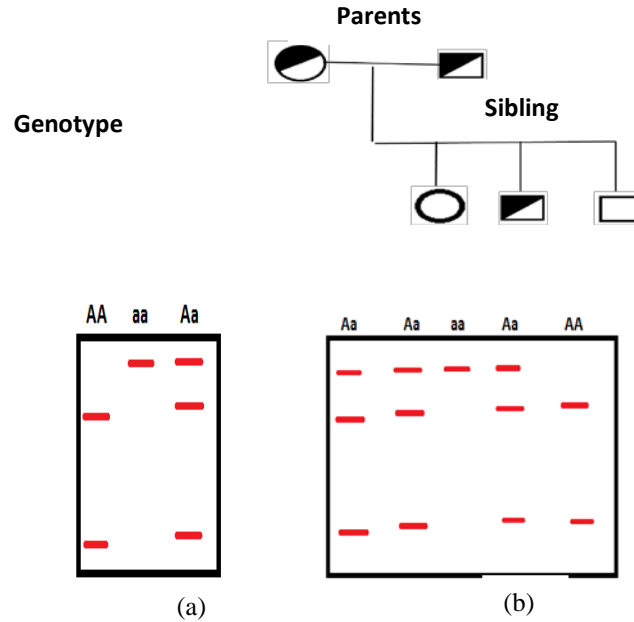


Figure 6-5.3(a).1.1 Restriction fragment length polymorphisms (RFLP): (a) shows the RFLP pattern homozygous and heterozygous parents.(b) The RFLP pattern of offspring which can be compared with parents to develop a RFLP map.

### Limitations of RFLP

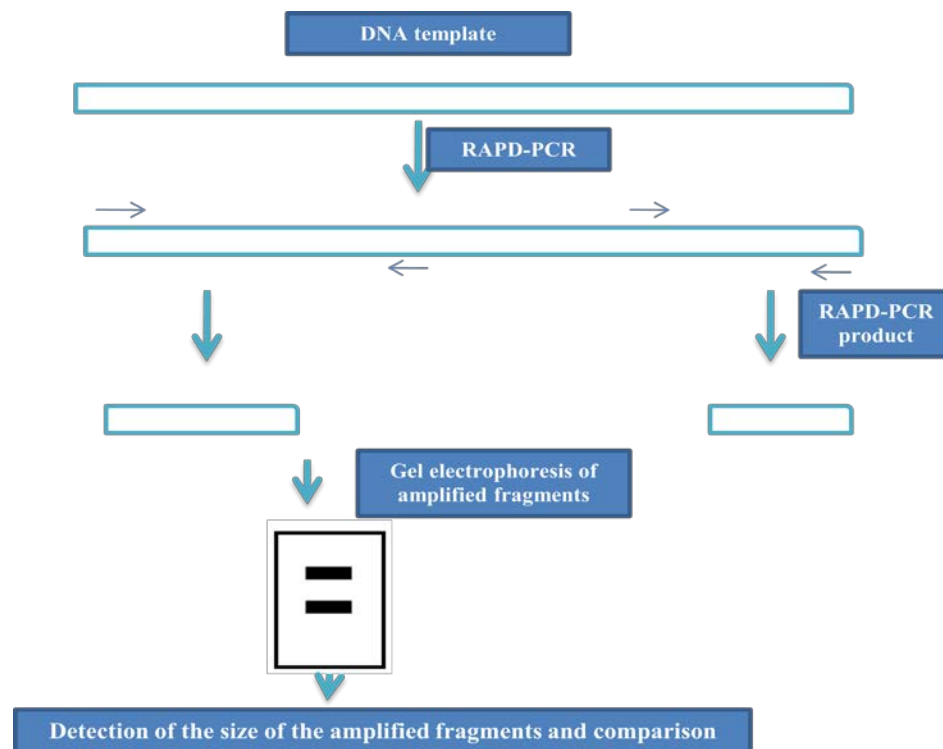
- Requires high quantity and quality of DNA
- Low polymorphism
- Requires development of specific probe libraries
- Requires radioactively labeled probes

Laborious and time consuming

### 6-5.3(a).1.2 Random amplification of polymorphic DNA (RAPD)

- It is the first molecular marker for PCR based genetic-mapping as well as DNA-fingerprinting.
- In the field of genetic mapping PCR-based markers are considered as second generation molecular markers.
- RAPD markers are short of length (approx 10 nucleotides only)
- RAPD needs only one primer instead of a set of primers for amplification.
- It does not require any prior information about the DNA sequence of the desired organism.

- Polymorphism i.e. relatively variable DNA sequences between different species, occurs due to mutation or rearrangements either at or in between the primer binding sites.
- PCR products are represented by electrophoresis and visualized by ethidium bromide staining.
- It is a quick, simple and efficient technique as it does not involve blotting or hybridization steps.
- Requires only small amount of DNA (10 ng/reaction) and the process can be automated.
- Primers are non species specific and can be universal.
- RAPD products can be cloned, sequenced and converted to other types of markers (SCAR, SNP)



**Figure 6-5.3(a).1.2 Steps involved in RAPD mapping: The DNA templates are used for PCR with RAPD random primers. After completion of the PCR cycles the product was visualized on agarose gel electrophoresis.**



## Limitations

- RAPD provides predominantly dominant markers.
- It is not possible to identify that whether the amplified region originated from a DNA segment locus that is heterozygous or homozygous.
- In case of co-dominant, RAPD markers show different-sized DNA fragments after amplification from the same locus and detected very rarely.
- It is fully dependent on PCR reaction reagents as well as cycle conditions.

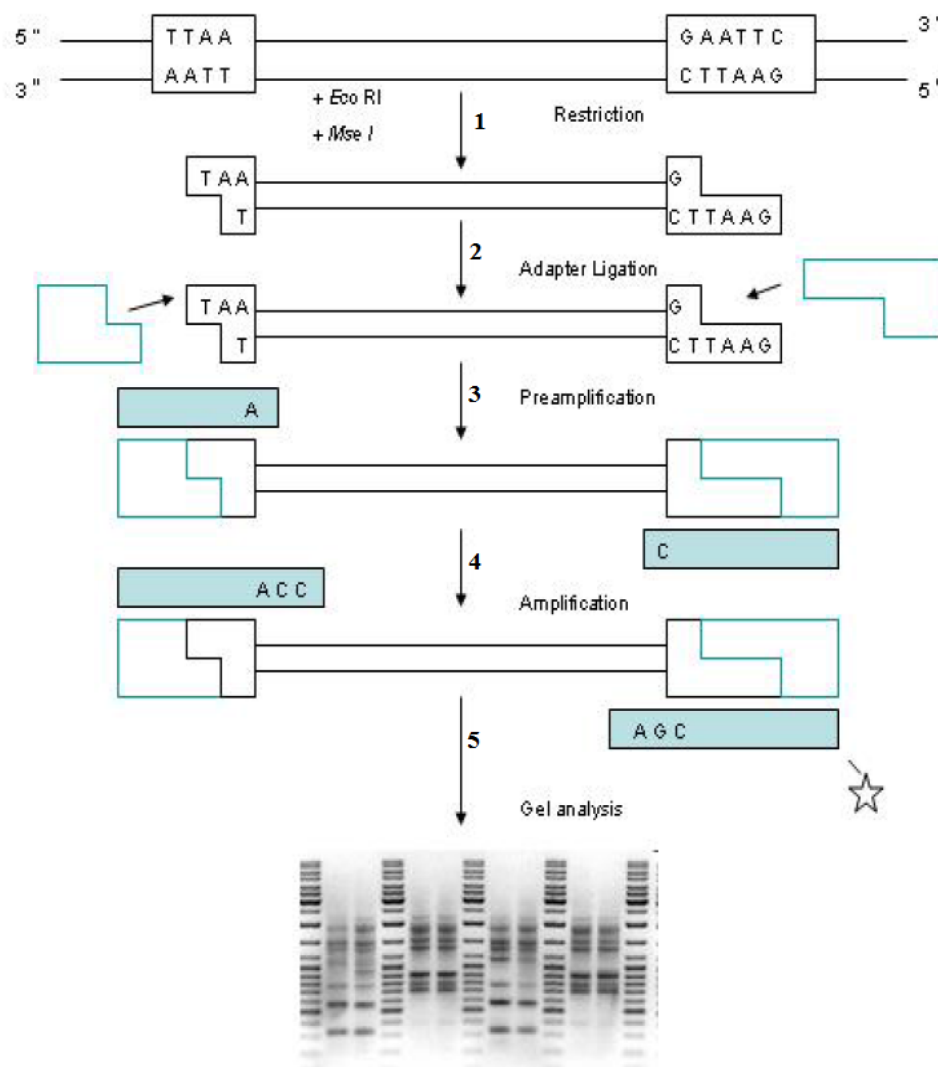
### 6-5.3(a).1.3 Amplified Fragment Length Polymorphism (AFLP)

- AFLP is also known as selective restriction fragment amplification (SRFA) as it involves selective PCR amplification of restriction fragments from a total double digest of genomic DNA under stringent condition.
- AFLP is more efficient than RFLP & RAPD in detection of polymorphism and has high reproducing ability.
- Most AFLPs are well known as dominant markers.
- The AFLP technique involves both the previous mapping basics such as restriction digestion of RFLP and PCR amplification of RAPD.
- The additional step in AFLP is ligating the fragments with adaptors of known sequence against which the primer has to be designed followed by PCR.
- One of the limitations of AFLP is that it is too complicated to be suitable for genotyping or MAS (Marker assisted selection).

AFLP comprises of three main steps (Figure 6-5.3(a).1.3) (Vos et al., 1995).

1. Cutting of genomic DNA with two restriction endonuclease enzymes simultaneously.
2. Ligation of adapters to the genomic DNA fragments to generate target sites for primer annealing and selective amplification of a subset of genomic restriction fragments by polymerase chain reaction (PCR). PCR amplification involves two consecutive reactions. In first PCR reaction (pre-amplification) DNA fragments are amplified with two AFLP primers, each having one selective nucleotide. In

- the second PCR reaction (selective amplification) the products are diluted and used as templates for the amplification using two AFLP primers, each having three selective nucleotides.
3. Separation and detection of the amplified fragments DNA using denaturing polyacrylamide gel electrophoresis and silver staining respectively.



**Figure 6-5.3(a).1.3 Steps involved in AFLP** (1) Genomic DNA is digested with two restriction enzymes (in this case *EcoRI* and *MseI*) (2) Ligation of specific adapters to the digested fragments at both ends (3) & (4) PCR is performed in two steps first is pre-amplification and followed by amplification. (5) This is to visualize the result of AFLP in agarose gel and detection of bands through silver staining or by using labeled primer (i.e. radioisotope or a fluorochrome). It is indicated by a star in the figure. (Adapted and modified from: Gabriel Romero et al 2009, *Genetic fingerprinting: Advancing the frontiers of crop biology research*)

#### 6-5.3(a).1.4 Single Nucleotide Polymorphisms (SNPs)

- Polymorphisms resulting from point mutations are the most abundant polymorphism in organisms. They can give rise to different alleles comprising alternative bases at a given nucleotide position within a locus.
- Advent of gene chip technology has improved the ability to genotype these single nucleotide polymorphisms (SNPs) in a large number of samples. These SNPs have emerged as a central point in the development of molecular markers.
- SNPs marker development can be automated, and have the power to reveal hidden polymorphism not detected with other markers and methods.
- There is very high number of SNPs in each genome which may help to generate RFLPs if there is any restriction site in the sequence.
- At least 1.42 million SNPs are found in the human genome.
- Considering SNP as a single alphabet spelling mistake (A, T, C, and G) in a word (stretch of DNA) theoretically, a SNP locus can have up to four variant alleles.
- However in the real scenario, most SNPs are usually restricted to one of two alleles (most often either the two pyrimidines C/T or the two purines A/G) and thus considered predominantly as bi-allelic. They are getting inherited as co-dominant markers. Although the polymorphism information content (PIC), i.e. the measure of informativeness of a genetic marker to detect polymorphism, of SNPs is not as high as multi-allele microsatellites, this shortcoming is compensated by their remarkable abundance in the genome.

Methods available for SNP genotyping include:

1. Traditional methods as listed below suitable for small laboratories limited by budget and labor constraints
  - Direct sequencing,
  - Allele-specific oligonucleotide (aso) (Malmgren et al., 1996),
  - Single strand conformational polymorphism assays (sscp) (Suzuki et al., 1990),
  - Single base sequencing (Cotton, 1993),
  - Denaturing gradient gel electrophoresis (dgge) (Cariello et al., 1988) and

- Ligation chain reaction (lcr), (Kalin et al., 1992).
- 2. Modern methods that can undertake large-scale analysis of SNP markers which are expensive and require sophisticated equipment,
- Matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometry (Ross et al., 1998, Storm et al., 2003),
- Pyrosequencing (Ahmadian et al., 2000),
- Taqman allelic discrimination (Li et al., 2004),
- Real-time (quantitative) PCR (Nurmi et al., 2001), and
- Microarray or gene chips (Hacia et al., 1999).

#### 6-5.3(a).1.5 Multi locus probe markers

The non-coding DNA, which makes up a large proportion of the genomes of higher organisms, contains regulatory elements like promoters and enhancers and repetitive elements (Turner et al., 1998) viz. mini- and microsatellite DNA. The term microsatellite was coined by Litt and Luty, and the term minisatellites was introduced by Jeffrey. They exist as multiple allelic forms in a population, have high level of heterozygosity, follow Mendelian inheritance, and behave as dominant fingerprinting markers and co dominant STMS (sequence tagged microsatellites) markers. Both micro- and minisatellites form an ideal marker system to create complex banding patterns through simultaneous detection of multiple DNA loci.

#### 6-5.3(a).1.5(i) Minisatellites

Minisatellites were first found in the human genome and described as hypervariable tandem repeats with a monomer repeat length of about 10–100 bp (more than the microsatellite), that is repeated several times (less than the microsatellite). This number of repeat units in loci varies between genotypes and is referred to as variable number of tandem repeats (VNTRs) or hypervariable regions (HVRs). Although both mini- and micro-satellites occur throughout the eukaryotic genome, the former tend to be

concentrated in the telomere regions and in sites associated with a high frequency of recombination (Bruford & Wayne, 1993; Nicholas, 1996).

#### 6-5.3(a).1.5(ii) Microsatellite

Microsatellite also known as short tandem repeats (STRs) or simple sequence repeats (SSR) are short repeating nucleotides (2-6 nts). In plants di-, tri-, and tetra- nucleotide repeats like (GT)<sub>n</sub>, (AAT)<sub>n</sub>, and (GATA)<sub>n</sub> are found widely distributed. In fishes microsatellites have been estimated to occur once in every 10 kb (Wright, 1993) of DNA length. Copy number of these repeats varies from individual to individual and forms the basis of polymorphism. One of the illustrative examples of microsatellites within coding regions are those causing genetic diseases in humans, such as the CAG repeats that encode polyglutamine tract, resulting in mental retardation. Microsatellite markers are highly polymorphic, and occur in large numbers. It is co-dominant and has ability to distinguish multiple alleles in a plant. Two conserved and locus-specific PCR primers flanking each microsatellite repeat are used for PCR amplification.

Several detection methods have been developed based on primer labeling, gel utility, staining, and detection equipment. Another way to detect SSR products is carried out on a DNA sequencer that detects fluorescent dye labels by laser excitation.

#### 6-5.3(a).1.6 Inter Simple Sequence Repeat (ISSR)

- As the name suggests “inter simple sequence repeat” involves amplification of DNA region that is located in between the two identical microsatellite repeats that are oriented in opposite direction. Briefly ISSR analysis helps in the understanding of organization, frequency and levels of polymorphism of different SSR's in a genome (Reddy 2002). ISSR is a PCR based method where 16–25 bp long microsatellites are used as primers in a single primer PCR reaction to target multiple genomic loci for the amplification of mainly the inter- SSR sequences of variable sizes.
- The primers used are either unanchored (Gupta et al., 1994; Meyer et al., 1993) or anchored at 5' or 3' end with 1-4 degenerate bases extended into the flanking sequences (Zietkiewicz et al., 1994).

- ISSR associates most of the advantages of AFLP and microsatellite analysis with the universality of RAPD. However as longer primers are used as compared to RAPD primers (10- mers) deployment of high annealing temperature (45–60°C) provides higher reproducibility and stringency.
- ISSRs mostly segregate as dominant markers following simple Mendelian inheritance (Gupta et al., 1994) and in few cases as co-dominant markers. The latter enabling distinction between homozygotes and heterozygotes (Wu et al., 1994; Akagi et al., 1996).

The major areas of the application of ISSR are as below,

- i. Genomic fingerprinting
- ii. Genetic diversity and phylogenetic analysis
- iii. Genome mapping
- iv. Determining SSR motif frequency
- v. Gene tagging and use in marker assisted selection
- vi. Evolutionary biology.

There are many other markers which can be used for mapping in plants. Few of them explained above and the comparative study of the five most commonly used DNA markers are give in table 6-5.3(a).1.

**Table 6-5.3(a).1 Comparative study of the five most commonly used DNA markers for genetic mapping in plants**

<b>Feature and description</b>	<b>RFLP</b>	<b>RAPD</b>	<b>AFLP</b>	<b>SSR</b>	<b>SNP</b>
Genomic abundance	High	High	High	Moderate to High	Very High
Genomic coverage	Low copy coding region	Whole genome	Whole genome	Whole genome	Whole genome
Expression/inheritance	Co-	Dominant	Dominant	Co-	Co-dominant

	dominant		/Co-dominant	dominant	
Number of loci	Small (<1,000)	Small (<1,000)	Moderate (1,000s)	High (1,000s – 10,000s)	Very high (>100,000)
Level of polymorphism	Moderate	High	High	High	High
Type of polymorphism	Single base changes, indels	Single base changes, indels	Single base changes, indels	Changes in length of repeats	Single base changes, indels
Type of probe/primers	Low copy DNA or cDNA clones	10 bp random nucleotides	Specific sequence	Specific sequence	Allele-specific PCR primers
Cloning and/or sequencing	Yes	No	No	Yes	Yes
PCR-based	Usually no	Yes	Yes	Yes	Yes
Radioactive detection	Usually yes	No	Yes or No	Usually no	No
Reproducibility/reliability	High	Low	High	High	High
Effective multiplex ratio	Low	Moderate	High	High	Moderate to High
Marker index	Low	Moderate	Moderate to High	High	Moderate
Genotyping throughput	Low	Low	High	High	High
Amount of DNA required	Large (5 – 50 µg)	Small (0.01 – 0.1 µg)	Moderate (0.5 – 1.0 µg)	Small (0.05 – 0.12 µg)	Small ( $\geq 0.05$ µg)
Quality of DNA	High	Moderate	Moderate to	High	High

required			High		
Technically demanding	High	Low	Moderate	Low	Low
Time demanding	High	Low	Moderate	Low	Low
Ease of use	Not easy	Easy	Moderate	Easy	Easy
Ease of automation	Low	Moderate	Moderate to High	High	High
Development/start-up cost	Moderate to High	Low	Moderate	Moderate to High	High
Cost per analysis	High	Low	Moderate	Low	Low
Number of polymorphic loci per analysis	1.0 – 3.0	1.5 – 5.0	20 – 100	1.0 – 3.0	1.0
Primary application	Genetics	Diversity	Diversity and genetics	All purposes	All purposes

**Table6-5.3(a).1 Comparison between most widely used DNA marker; Adapted from Collard et al. (2005), Semagn et al. (2006a), Xu (2010), and others also given in “Plant Breeding from Laboratories to Fields”, Chapter 3(Molecular Markers and Marker-Assisted Breeding in Plants) Page 52 & 53.**

### 6-5.3(b)Physical Mapping

A physical map represents a partial ordering of specific DNA sequences by their position on a chromosome. It has been developed to address the limitations of the maps generated by genetic techniques listed below-

- Less accuracy of genetic maps
- Resolution of genetic maps depends upon the number of cross-overs to be scored.

There are several types of physical maps including cytological maps, restriction maps, STS content maps, radiation hybrid maps, ordered clone collections (i.e., “contig maps”), and the DNA sequence of an entire chromosome, constructed by a wide range of experimental methods without having any complete sequence information. Various techniques employed for the generation of physical maps are described below.



- 1) Restriction Mapping
- 2) Fluorescent *in situ* hybridization (FISH)
- 3) Sequence tagged site (STS) mapping

#### 6-5.3(b).1 Restriction Mapping

Restriction mapping helps in obtaining structural information of an unknown piece of DNA. A restriction map is generated by digesting DNA with a series of restriction enzymes followed by separation of the resultant DNA fragments by agarose gel electrophoresis. The distance between restriction enzyme sites is determined from the fragmentation pattern in the gel.

In practice a DNA molecule is first digested with a known restriction enzyme (say *EcoRI*) and the sizes of the resulting fragments are measured by agarose gel electrophoresis. Then this is repeated with another enzyme (say *BamHI*) that identifies a different consensus sequence. From the gel patterns the numbers of restriction sites for each enzyme are calculated. To determine the relative positions of the restriction sites w.r.t each enzyme, in yet another step DNA molecule is digested with both enzymes together.

In case of problems arising due to occurrence of more than one restriction sites of the second enzyme in a fragment generated by the first enzyme the solution is to undertake partial restriction using less enzyme or short reaction time whenever required. Finally the positions of the restriction sites obtained through the above process is plotted on a bar representing a DNA segment to obtain a restriction map.

#### 6-5.3(b).2 Fluorescent *in situ* hybridization (FISH)

The principle of this method is described in detail in Module 3-Lecture 1. It was first employed with metaphase chromosomes to locate the position of a marker on a chromosome or extended DNA molecule using fluorescent probes leading to low-resolution which was insufficient for the construction of useful chromosome maps. Since 1995, various high resolution techniques have been developed for accurate mapping of chromosomes by altering the chromosomal nature as described below,

1. Mechanical stretching of chromosomes which can be done by altering the isolation method of metaphase chromosomes by centrifugation resulting in the stretching of the chromosomes several times to their normal length. FISH signals are mapped with improved resolution; markers 200-300 kb of length can be mapped by this method.

2. Another approach is to map the non-metaphase chromosomes (those in prophase and interphase), which are sufficiently condensed to be easily identified. Resolution obtained is 25 kb but there is a loss in the characteristic chromosome morphology for which preliminary map information is required to locate the position of markers on a small region of chromosome.

To overcome the above limitation, **fiber-FISH** has emerged as a high resolution method (>25 kb) which uses DNA prepared by gel stretching or molecular combing and can distinguish markers lying within 10 kb.

Gel stretching:

The chromosomal DNA is suspended in molten agarose and pipetted onto the microscope slide coated with restriction enzyme (inactive stage). With the cooling and solidification of agarose, the DNA molecules become stretched. When magnesium chloride is added, the magnesium ions activate the restriction enzyme which cuts these molecules. With the gradual coiling of molecules, the gaps representing the cuts can be visible under fluorescence microscope (Figure 6-5.3(b).2(i)).

Molecular combing:

In it, silicone-coated cover slip is dipped in DNA solution where these molecules get attached to it by their ends. With the withdrawal of cover slip at a rate of  $0.3 \text{ mm s}^{-1}$ , these DNA molecules get retained as an array (comb) of parallel molecules (Figure 6-5.3(b).2(ii)).

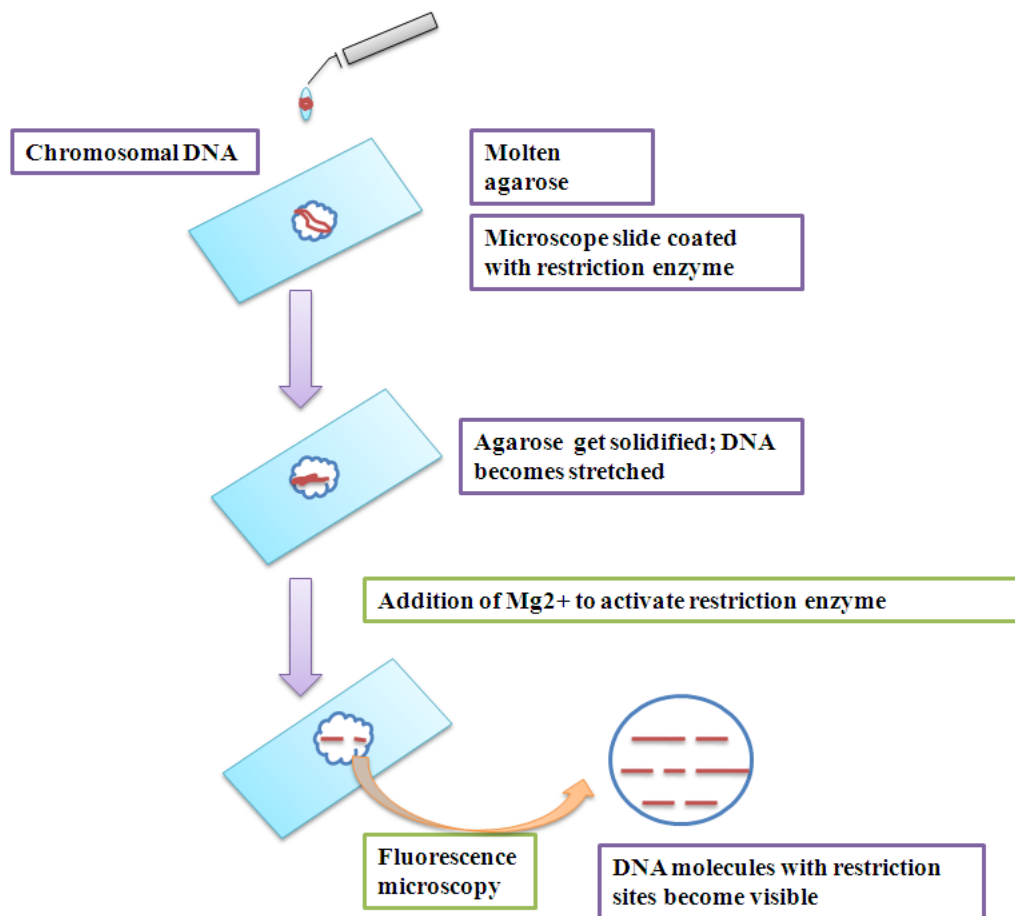


Figure 6-5.3(b).2(i) Gel stretching

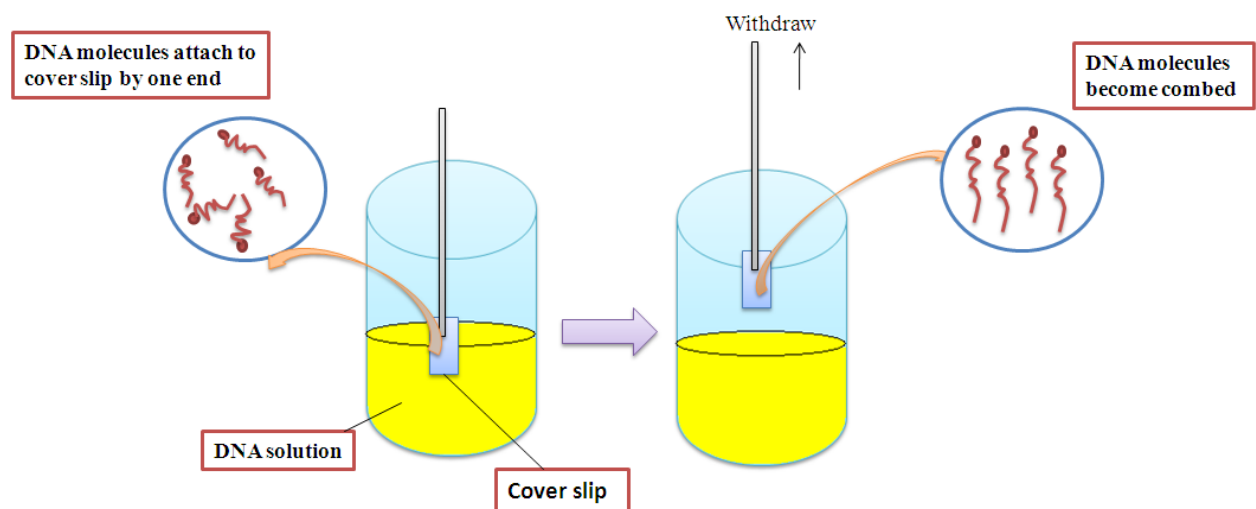
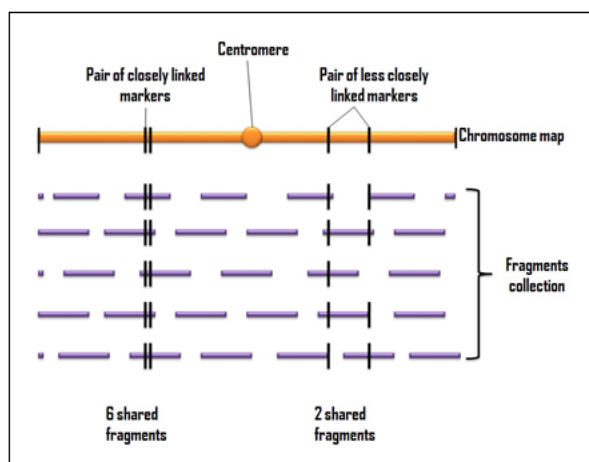


Figure 6-5.3(b).2(ii) Molecular combing

### 6-5.3(b).3 Sequence Tagged Site (STS) Mapping

Both restriction mapping and FISH face one or more disadvantages. Restriction mapping is easy and rapid technique providing detailed information, but it is not applicable for large genomes. Although FISH can be applied to larger genomes, and fiber-FISH can provide high-resolution data, it is a technically challenging method to implement. STS can overcome many of these limitations and give a detailed map with comparative ease.

STS is an easily recognizable short DNA sequence, typically 100-500 bp in length, generated by PCR using primers obtained from already known sequences. STS (sequence tagged site) markers are used for detailed mapping of large genomes. This method involves collection of various overlapping DNA fragments obtained from either single chromosome or complete genome. The distance between the markers on these fragments determines which fragment contains which STSs (Figure 6-5.3(b).3.1). Closer the markers in the genome, greater will be the chance of occurrence of two STSs on the same fragment.



**Figure 6-5.3(b).3.1 STS mapping representing the collection of various fragments that span the entire length of a chromosome, with each point on the chromosome present in an average of five fragments. (Adapted and modified form Brown TA. 2002.**

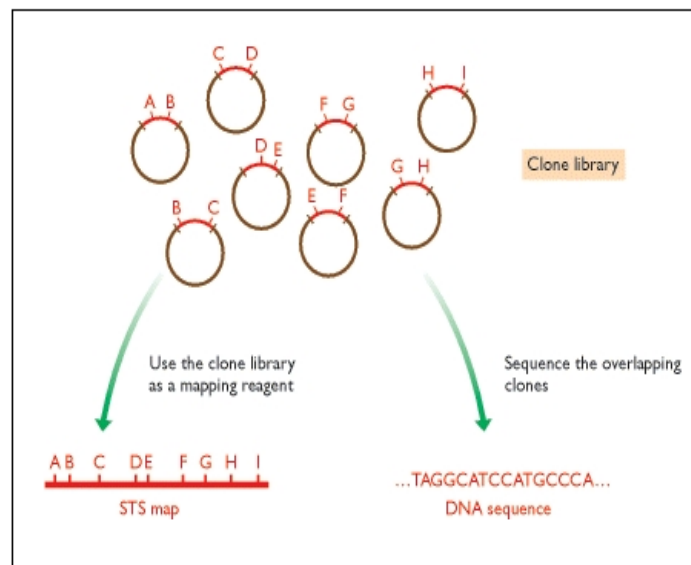
*Genomes. 2<sup>nd</sup> ed. Oxford: Wiley-Liss)*

Various sequences such as ESTs, SSLPs and random genomic sequences can be used as a source for STS mapping. Expressed sequence tags (ESTs) are short sequences obtained by analysis of complementary DNA (cDNA) clones. They represent the sequences of the genes being expressed in a cell from which mRNAs are derived. It is commonly used as

STS as it comes from a unique gene rather than from a member of a gene family having same or similar gene sequences. Random genomic sequences may be obtained by downloading sequences deposited in the databases or by sequencing random fragments of cloned DNA.

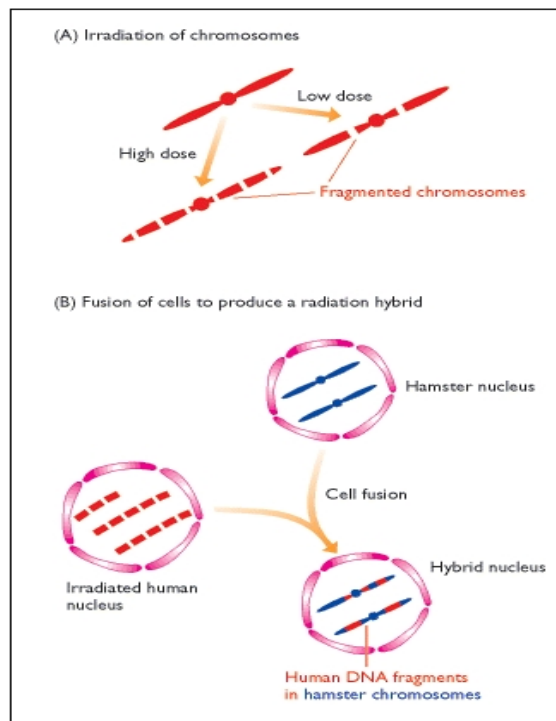
Assembly of fragments (often referred as mapping reagent) can then be performed in two ways- as a clone library or as radiation hybrids.

A clone library can be used as mapping reagent in STS analysis having an average size of several hundred kb represented in Figure 6-5.3(b).3.2. Chromosome-specific library can be constructed by separating the chromosomes by flow cytometry.



**Figure 6-5.3(b).3.2. Clone libraries used in the construction of STS maps and also as a source of DNA to be sequenced.**  
(Source: Brown TA. 2002. *Genomes*. 2<sup>nd</sup> ed. Oxford: Wiley-Liss)

A radiation hybrid can also be used as a mapping reagent in STS mapping which refers to a hamster cell line carrying a relatively small DNA fragment from the genome of a second organism (such as human). For example, formation of radiation hybrid between human cell and hamster cells shown in Figure 6-5.3(b).3.3. After fusion of the two cells, the presence of markers in human DNA within the hybrid cell line is analyzed by hybridization or PCR methods. Same radiation hybrid will result if the two markers are close together thereby having greater probability to be on the same DNA fragment.



**Figure 6-5.3(b).3.3. Radiation hybrids(a) Irradiation of human cells: formation of smaller fragments by low X-ray doses and larger fragments by higher doses (b) Formation of a radiation hybrid by fusion between an irradiated human cell and an untreated hamster cell.**

(Source: Brown TA. 2002. *Genomes*. 2<sup>nd</sup> ed. Oxford: Wiley-Liss)

## Bibliography

Alonso *et al.* (2006) Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in *Arabidopsis*. *Nat. Rev. Genet.* 1-12

Altmann T (1999). A complete BAC-based physical map of the *Arabidopsis thaliana* genome. *Nature Genetics* 1999 Jul; 22(3):271-5.

Bortiriet *al.* (2006) Advanced in maize genomics: the emergence of positional cloning. *Curr. Opin. Plant Biol.* 9. 164-171

Brown TA. (2002). *Genomes*. 2<sup>nd</sup> ed. Oxford: Wiley-Liss.

Bruford, M.W. & Wayne, R.K. (1993). Microsatellites and their application to population genetic studies. *Curr. Op. in Gen. and Dev.* 3, 939-943.

Fujii J., Otsu K., Zorzato S. De Leon S., Khanna V.K., Weiler J.E., O'Brien P.J., MacLennan D.H. (1991). "Identification of a mutation in porcine ryanodine receptor associated with malignant hyperthermia". *Science* 253: 448.

Huang *et al.* (2005) Comparative genomics enabled the isolation of the *R3a* late blight resistance gene in potato. *Plant J.* 42, 251-261

Huang N., Cogburn L.A., Agarwal S.K., Marks H.L., Burnside, J. (1993). Over expression of a truncated growth hormone receptor in the sex-linked dwarf chicken: evidence for a splice mutation. *Molecular Endocrinology* 7: 1391.

Janderet *al.* (2002) Arabidopsis map-based cloning in the post-genome era. *Plant physiol.* 129, 440-450

Jeffrey, A. J., Wilson, V. and Thein, S. L. (1985). *Nature*, 314, 67–73.

Keller *et al.* (2005) Map-based isolation of disease resistance genes from bread wheat: cloning in a surprise genome. *Genet. Res.* 85, 93-100

King & Motulsky (2002). *Science* 298, 2342-2343 Rosenberg et al. (2002). *Science* 298, 2381-2385.

Klungland H., Vage D.I., Raya L.G., Adalsteinsson S., Lien S. (1995). "The role of melanocyte-stimulating hormone (MSH) receptor in bovine coat color determination". *Mammalian Genome* 6: 636.

Litt, M. and Luty, J. A., Am. J. (1989), *Hum. Genet.* 44, 397–401.

Lukowitzet *al.* (2000) Positional cloning in Arabidopsis. Why it feels good to have a genome initiate working for you. *Plant Physiol.* 123, 795-805

Maria Sicaet *al.* (2005) ISSR markers show differentiation among Italian populations of *Asparagus acutifolius* L, *BMNC Genetics* . 6:17

Martin *et al.* (1993) Map-based cloning of a protein *kinase* gene conferring disease resistance in tomato. *Science* 262, 1432-1436

Mauricio (2001) Mapping quantitative trait loci plants: uses and caveats for evolutionary biology. *Nat. Rev. Genet.* 2, 370-381

Montgomery G.W., Lord E.A., Penty J.M., Dodds K.G., Broad T.E., Cambridge L.M., Sunden S.L.F., Stone R.T., Crawford A.M. (1994). “The Booroola fecundity (FecB) gene maps to sheep chromosome 6”. *Genomics* 22 148.

Morton N.E., (1955). “Sequential tests for the detection of linkage”. *American journal of human genetics* 7: 277.

Peters *et al.* (2003) Forward genetics and map-based cloning approaches. *Trends Plant Sci.* 8, 484-491

Reece R.J. (2000). *Analysis of Genes and Genomes*. John Wiley & Sons, U.K.

Tai *et al.* (1999) Expression of *Bs2* pepper gene confers resistance to bacterial spot disease in tomato. *PNAS* 96, 14153-14158

Turner, P.C., McLennan, A.G., Bates, A.D. & White, M.R.H. (1998). *Instant notes in molecular biology*. BIOS Scientific Publishers Limited, Oxford, UK.

Vaiman D., Koutita O., Oustry A., Elsen J.M., Manfredi E., Fellous M., Cribiu E.P. (1996). “Genetic mapping of the autosomal region involved in XX sex-reversal and horn development in goats”. *Mammalian Genome* 7:133.

Xiong M, Chen HJ, Prade RA, Wang Y, Griffith J, Timberlaken WE, Arnold J. (1996). On the Consistency of a Physical Mapping Method to Reconstruct a Chromosome *in Vitro*. *Genetics Society of America*.



## **MODULE 6- LECTURE 6**

# **MARKER-ASSISTED SELECTION FOR PLANT BREEDING AND LIVESTOCK IMPROVEMENT**

### **6-6.1 Introduction**

Traditional breeders have very effectively manipulated the genomes of both crops and livestock species, exploiting the existing natural variations within a species, a breed, and a population. Early breeders lacked knowledge of molecular genetics or selection markers and enhanced the production traits in their herds through selection of superior individuals as progenitors for next generations. It required dedication and hard work of combining phenotypic records of individual performance with genealogical information often for many years for achieving enhanced “breeding values”.

Increased milk production of 110 kg per animal per year in Holstein dairy cattle or the attainment of 50% decreased feed conversion (the kilograms of feed required to produce a kilogram of pork,) in pig production between the 1960s and 2005 are few examples to illustrate the power of traditional breeding. However, the efficiency of traditional methods decreases in cases where there is difficulty to measure traits, low heritability, or cannot be correctly, quickly and inexpensively measured in a large number of animals (Eggen 2012). These difficult-to-measure traits including fertility, longevity, feed efficiency, and disease resistance are often critically important and the selection of these traits must be achieved through genomic means.

### **6-6.2 Marker Assisted Selection**

The use of DNA markers to target the loci to assist phenotypic screening is known as Marker Assisted Selection (MAS). MAS help to make selection decisions before phenotypes are available. A broader terminology is Marker Assisted Breeding (MAB) which includes apart from MAS, several modern breeding strategies like marker-assisted recurrent selection (MARS), marker-assisted backcrossing (MABC), genome-wide selection (GWS) or genomic selection (GS) etc (Ribaut et al., 2010). MAB is the

application of molecular markers along with linkage maps and genomics to alter and improve plant or animal traits based on genotypic assays.

An ideal marker is the one which allows inheritance to be followed across generations. Recent advances in molecular genetics offer methods to dissect the genetic variability of complex traits into quantitative trait loci (QTL). Generally a marker is associated to a QTL. The best scenario is when the marker and QTL are in the same locus. For practical purposes the marker and QTL should be closely located so that recombination between marker and QTL is rare.

Most traits of economic importance are complex quantitative traits controlled by many genes. However, some of these genes might have a larger effect than the others and are called as major genes located at QTL. QTL strictly applies to genes of any effect, however in practice it applies only to major genes, having effects large enough for detection and mapping the genome. The pattern of inheritance of such genes may assist in selection.

Thus marker-assisted selection helps in the accurate selection of specific DNA variations that have been associated with a measurable difference or effect on complex traits. It is critical to understand that markers for complex traits like milk yield etc. are associated with only one of the many genes (major gene) that contribute towards that trait. The presence or absence of the numerous other “unmarked” genes both minor and/or other major (if any) and the production environment will ultimately determine whether an animal actually express the trait at the desired intensity. Therefore MAS is considered as an additional tool, and not as a replacement for, traditional selection techniques.

The potential benefits from MAS are maximum for the following traits which

- 1) Have low heritability,
- (2) Are difficult or expensive to measure,
- 3) Are not measured until the animal had already contributed to next generation,
- 4) Are currently not selected because they are not routinely measured,
- 5) Are genetically correlated with a trait that the breeders do not want to increase.

Current **MAS schemes** consist of the following steps (Meuwissen 2003),

1) Determine the biggest, statistically significant QTL(s) in a genome wide scan for QTL. There are two approaches to locate a QTL (i) Genome scan and (ii) Candidate gene approach.

2) Select for these big QTL next to selecting for polygenes (the remaining often being unidentified smaller genes).

### **6-6.3 MAS for Genetic Improvement of Crops and Livestock**

Markers must be tightly-linked to target loci. Ideally markers must be more than 5 cM (centimorgan) away from a gene or QTL for its linkage confirmation.

Using a pair of flanking markers greatly improve the reliability but increases time and cost.

The markers used for selection of a trait must be polymorphic to distinguish the banding pattern among same or different species.

It is important to locate all markers and protein coding genes in the chromosomes.

For detecting QTL based on genetic markers, development of genetic maps of the species of interest is required.

Development of PCR based techniques and microsatellite markers, increases the efficiency of the mapping procedure in all living organisms.

### 6-6.3.1 Prerequisites for a successful MAS programme

The crucial requirements for making a MAS programme successful are summarized below (Xu2003).

**i. High throughput DNA extraction.** Extraction of genomic DNA from plants is a limitation associated with sample processing for MAS. Various methods involving liquid nitrogen, phenol-chloroform, detergents, polyvinylchloride and salts (to remove polysaccharides and polyphenols) are available which are tedious and labor-intensive. A rapid, novel and automated genomic DNA extraction method from leaves of the plant is required that is compatible for PCR (polymerase chain reaction) applications.

**ii. Genetic markers.** Both DNA and other types of markers (protein, morphological, cytological) can be used in MAS programs. DNA markers are advantageous as they are polymorphic, distributed in large numbers throughout the genome, their presence or absence unaffected by environment and do not directly affect the phenotype. The commonly used DNA markers like RFLP, SSR, RAPD, AFLP and SNP, have been discussed in Module 6 Lecture 5. Overall for successful MAS programme a marker should have the following attributes,

- Easy to detect
- Have high rate of polymorphism
- Co-dominance
- Occurrence throughout the genome
- Amount of DNA required be small
- Low/null interaction among the markers for use of many at the same time in a segregating population.

**iii. Genetic maps.** Linkage maps help in detecting marker-trait associations and in choosing markers for a MAS programme. After a marker is found to be associated with a trait in a given population, a dense molecular marker map from a standard reference population will help in the identification of other markers that are closer to or that flank, the target gene.

**iv. Knowledge of associations between molecular markers and traits of interest.**

The knowledge of markers associated with traits relevant to a breeding program is the most important ingredient for MAS which is obtained from QTL or gene mapping studies, classical mutant analysis, bulked segregant analysis (Michelmore et al., 1991) etc.

**v. Data management system.** The number of samples handled in a MAS program is very large, where each sample is evaluated for multiple markers. In such a situation data handling and processing becomes very important. An efficient system for labeling, storing, retrieving, and analyzing large data sets, and producing reports useful to the breeder should be put in place on priority basis.

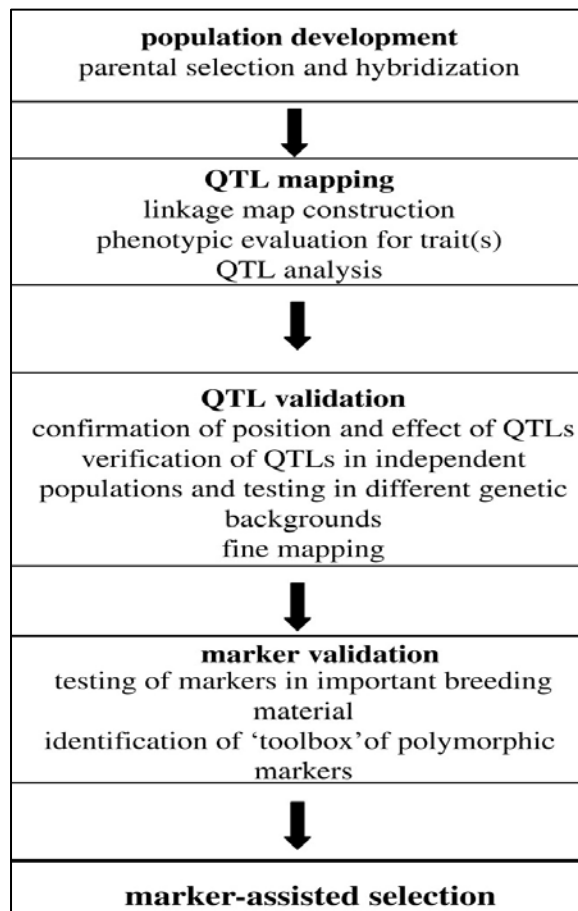


Figure6-6.3.1.Marker assisted selection (Adapted from Trans. R. Soc. B 2008;363:557-572)

## 6-6.4 Marker development and MAS

The overview of marker development as suggested by **Collard and Mackill (2008)** is presented in figure 6-6.3.1

### 6-6.4.1 Steps in QTL Mapping

The fundamental approach for QTL mapping is illustrated in figure 6-6.4.1 by taking into consideration the hair density (i.e. trichomes) that are present on a plant leaf.

Step1. Inbred parental lines having different trichome density (i.e. Dense vs. Sparse) are crossed. The  $F_1$  population thus obtained has intermediate density of trichomes.

Step2. It is then followed by selfing of  $F_1$  population to obtain  $F_2$  population.

Step3. Each  $F_2$  individual is selfed for 6 subsequent generations resulting in the formation of several recombinant inbred lines (RILs). Each RIL thus obtained is homozygous for a section of a parental chromosome. Scoring of these RIL's is done for several genetic markers and the trichome density phenotype.

Step4. The arrow represents a section of chromosome derived from the parent having sparse trichome density. Leaves of all the individuals that have inherited this section also have sparse trichome density, demonstrating that this chromosomal region probably contains a QTL for this trait.

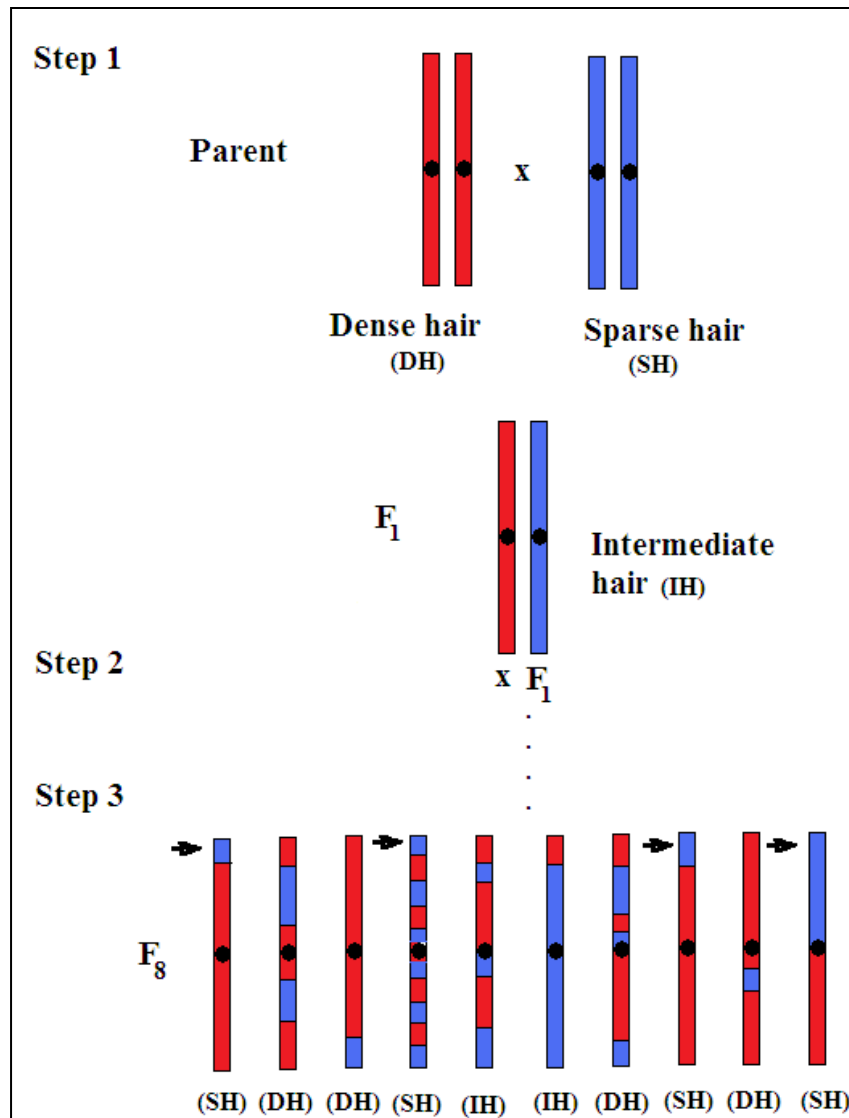


Figure 6-6.4.1. A schematic representation of QTL mapping. (Adapted from Mauricio, R. *Nature Reviews Genetics* (2001) 2: 370)

### 6-6.4.2 Relevance of QTL Mapping For MAS

QTL mapping is the foundation of the development of markers for MAS. Factors that affect the accuracy of QTL mapping are population size and type, level of replication of phenotypic data, environmental effects and genotyping errors. Following steps are involved in QTL mapping:

- i) **Marker conversion** may be required to make the marker genotyping method technically simpler for the improvement of MAS reliability.
- ii) **QTL confirmation** tests the accuracy of result from the primary QTL mapping studies.
- iii) **QTL validation** talks about the verification of the QTL's effectiveness in different genetic backgrounds.
- iv) **Marker validation** is required to test the level of polymorphism of tightly-linked markers within a very short distance (approx 5-10 cM) covering a target locus and also to test the reliability of markers in phenotype prediction.

## 6-6.5 MAS Schemes in Plant Breeding

### 6-6.5.1 Marker assisted backcrossing

Holland (2004) has described three general levels of marker-assisted backcrossing namely foreground, recombinant and background selection (figure 6-6.5.1).

**Foreground selection** where markers are employed in combination with or for replacing the screening for target gene or QTL (Hospital & Charcosset, 1997). This helps in

- a) Screening traits that have laborious or time-consuming phenotypic procedures,
- b) To select for reproductive-stage traits in the seedling stage allowing the best plants to be identified for backcrossing,
- c) Selection of recessive alleles, which is difficult to achieve by conventional methods.



**Recombinant selection** involves selection of BC progeny with the target gene and involves the recombination events that occur between the target loci and linked flanking markers (e.g. less than 5 cM on either side). This helps in reducing the size of the donor chromosome fragment comprising the target locus (i.e. size of the introgression) to minimize linkage drag' (Hospital, 2005). In conventional breeding, the donor segment can remain very large till many BC generations (e.g. more than 10; Ribaut & Hoisington, 1998; Salina *et al.*, 2003). Due to rare possibility of occurrence of double recombination events on both sides of target loci "recombinant selection" can generally be performed utilizing at least two BC generations (Frisch *et al.*, 1999a).

**Background selection** involves selection of BC progeny having greatest part of recurrent parent (RP) genome by using background markers. Background markers are those unlinked to the target gene or QTL on all other chromosomes. These are also described as, markers that can be used to select against the donor genome. This accelerates RP recovery as compared to conventional backcrossing which takes a minimum of six BC generations to recover the RP the same can be achieved by BC<sub>4</sub>, BC<sub>3</sub> or even BC<sub>2</sub> (Visscher *et al.* 1996; Hospital & Charcosset 1997; Frisch *et al.* 1999a,b), thus saving two to four BC generations.

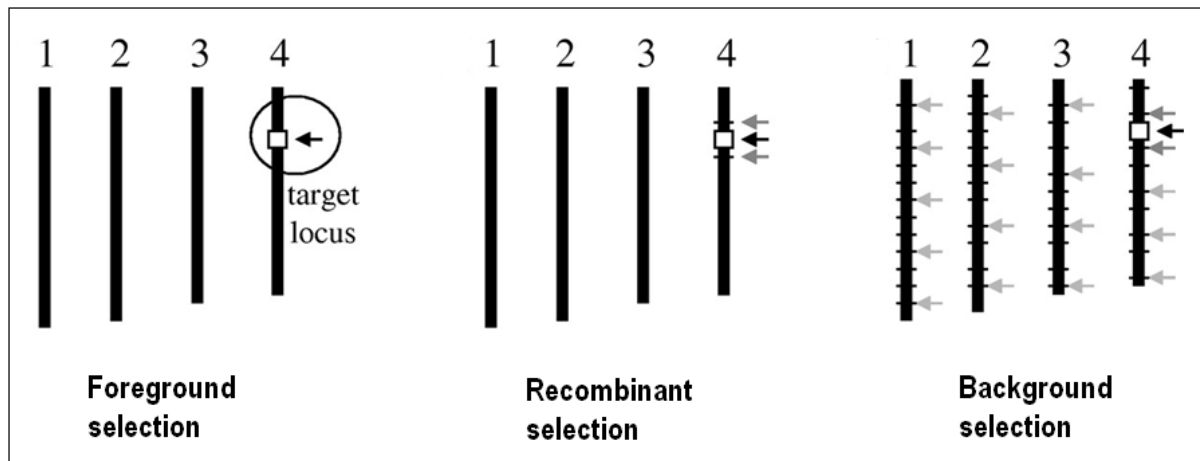


Figure 6-3.5.3.1 Recombinant and background selection. (Adapted from: *Trans. R. Soc. B* 2008;363:557-572)

**Marker Assisted Pyramiding** involves combining simultaneously multiple genes/QTLs together into a single genotype. Gene pyramiding is possible through conventional breeding but it is extremely difficult as individual animals must be phenotypically screened for all traits tested and nearly impossible at early generations. DNA markers can facilitate selection as they result in non-destructive assays. Also, with a single DNA sample, markers for multiple specific genes or QTLs can be easily verified without requiring phenotyping. These are mostly used for combining multiple disease resistance genes so as to develop durable disease resistance.

**Early generation marker assisted selection:** MAS provides a huge advantage that many plants with unwanted gene combinations, especially those lacking essential disease resistance traits and plant height, can be discarded in the early stage, so as to result in more efficient and cheap evaluation of other traits with fewer breeding lines at the later stages. However the major disadvantage of applying MAS at early generations is the high cost of genotyping a larger number of plants.

**Combined Approaches:** A combination of phenotypic screening and MAS approach can be utilized to maximize genetic gain (when some QTLs have been unidentified from QTL mapping) and to reduce the population sizes for such traits where marker genotyping is easier and cheaper than phenotypic screening.

### 6-6.6 Advantages of MAS

The efficiency and effectiveness for breeding may be greatly increased by MAS. The main advantages of MAS compared to conventional phenotypic selection are:

- It is simple and save time.
- Selection may be carried out at seedling stage.
- Single plants/animals may be selected with high reliability.
- Useful for traits that cannot be improved easily through phenotypic selection, either due to difficulty in measuring on young animals (i.e. before sexual reproduction), or due to low heritability.

### **6-6.7 Limitations of MAS**

1. It requires prior knowledge of gene alleles or markers associated with the traits of interest as well as the quantitative estimates of these associations in the specific population.
2. It explains only a limited part of the genetic differences between individuals.

Therefore in spite of the initial enthusiasm MAS created the outcome was not that promising. Currently using adequate genetic markers, one can follow the segregation of the entire genome in spite of a set of specific regions of interest, thereby, going from MAS towards genomic selection. Today parental relationships are no longer essential for explaining related performances in animals because, these can now be explained by the fact that animals are sharing identical chromosome regions using a dense set of SNPs, (Eggen 2012). The cost-effectiveness of whole-genome SNP panels for major livestock species and genomic selection tools is emerging as a paradigm shift in animal breeding. SNP panels for many species, including several agriculturally relevant species have been developed (Table. 6-6.7)

**Table6-6.7 List of whole-genome single nucleotide polymorphism (SNP) chips developed for important crop plants and livestock.**

<b>Species</b>	<b>Identification<sup>1</sup></b>	<b>Classification</b>	<b>Provider<sup>2</sup></b>	<b>Consortium</b>	<b>SNP, no.</b>
Potato	Potato	Public	Illumina	solCAP	8,303
Tomato	Tomato	Public	Illumina	solCAP	7,720
Apple	Apple	Public	Illumina	RosBREED	8,788
Peach	Peach	Public	Illumina	RosBREED	8,144
Cherry	Cherry	Public	Illumina	RosBREED	5,696
Maize	MaizeSNP50	Commercial	Illumina	Commercial	56,110
Rice	Rice44K	Commercial	Affymetrix	Commercial	44,100
Chicken	Chicken	Private: public sale	Illumina	Cobb Vantress- Hendrix- USDA	57,636
Cat	Feline	Private: public sale	Illumina	Morriss Animal Foundation	62,897
Horse	Equine	Private: public sale	Illumina	Neogen (Gene Seek)	65,157
Sheep	Ovine	Private: public sale	Illumina	AgResearch	5,409
Cattle	BovineHD	Commercial	Illumina	Various	777,962
Cattle	BovineSNP	Commercial	Illumina	Various	54,609
Cattle	BOS 1	Commercial	Affymetrix	Various	648,000
Sheep	OvineSNP50	Commercial	Illumina	Various	52,241
Cattle	BovineLD	Commercial	Illumina	Various	6,909
Pig	PorcineSNP60	Commercial	Illumina	Various	62,163
Dog	CanineHD	Commercial	Illumina	Various	173,662

1HD = high density; LD = low density, 2Illumina Inc., San Diego, CA; Affymetrix, Santa Clara, CA.(Source *Animal Frontiers* 2012, 2: 10-15)

3. In spite of various advantages the impact of MAS in crop improvement is low due to following reasons:

- Resources (equipment) not available.
- Markers may not be cost-effective.
- Accuracy of QTL mapping studies not high.
- QTL effects may depend on genetic background or may be influenced by environmental conditions.
- Lack of marker polymorphism in breeding material.
- Poor integration of molecular genetics and conventional breeding.

### **6-6.8 Future Challenges**

- Improved cost-efficiency- Optimization, simplification of methods and future innovation.
- Designing of effective and efficient MAS strategies.
- Greater integration between plant breeding and molecular genetics.
- Data management

## Bibliography

Collard B C, and Mackill D J *Phil. Trans. R. Soc. B* 2008; 363: 557-572.

Eggen, A. 2012. The development and application of genomic selection as a new breeding paradigm. *Animal Frontiers* 2(1): 10-15.

Frisch M, Bohn M, Melchinger A.E *Comparison of selection strategies for marker-assisted backcrossing of a gene. Crop Sci.* 39, 1999b 1295–1301.

Frisch M, Bohn M, Melchinger AE. *Minimum sample size and optimal positioning of flanking markers in marker-assisted backcrossing for transfer of a target gene. Crop Sci.* 39, 1999a 967–975.

Holland, J. B. 2004. Implementation of molecular markers for quantitative traits in breeding programs—challenges and opportunities. In *Proc. 4th Int. Crop Sci. Congress.*, Brisbane, Australia, 26 September—1 October.

Hospital F 2005 *Selection in backcross programmes. Phil. Trans. R. Soc. B.* 360, 1503–1511.

Hospital F, Charcosset A. 1997 Marker-assisted introgression of quantitative trait loci. *Genetics*. 147, 1469–1485.

Mauricio (2001) Mapping quantitative trait loci plants: uses and caveats for evolutionary biology. *Nat. Rev. Genet.* 2, 370-381

Meuwissen, T. Genomic selection: the future of marker assisted selection and animal breeding. *Proceedings of International workshop on Marker assisted selection: a fast track to increase genetic gain in plant and animal breeding? Session II: MAS in animals*, page 54, held on 17-18 October 2003 in Turin, Italy  
<http://www.fao.org/biotech/docs/Meuwissen.pdf>

Michelmore, R.W., I. Paran, and R.V. Kesseli. 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers

in specific genomic regions by using segregating populations. *Proc. Natl. Acad. Sci. USA.* 88:9828-9832.

Ribaut J.-M, Hoisington D 1998 *Marker-assisted selection: new tools and strategies. Trends Plant Sci.* 3, 236–239.

Ribaut, J.M., M.C. de Vicente, and X. Delannay. 2010. Molecular breeding in developing countries: challenges and perspectives. *Current Opinion in Plant Biology* 13: 1-6.

Salina E, Dobrovolskaya O, Efremova T, Leonova I, Roder M.S 2003 *Microsatellite monitoring of recombination around the Vrn-B1 locus of wheat during early backcross breeding. Plant Breed.* 122, 116–119.

Visscher P.M, Haley C.S, Thompson R 1996 Marker-assisted introgression in backcross breeding programs. *Genetics.* 144, 1923–1932.

Xu, Y. 2003. Developing marker-assisted selection strategies for breeding hybrid rice. *Plant Breeding Reviews* 23:73-174.

**Website:**

[http://www.agrireseau.qc.ca/bovinsboucherie/documents/Marker\\_Assisted\\_Selection\\_in\\_Beef\\_Cattle.pdf](http://www.agrireseau.qc.ca/bovinsboucherie/documents/Marker_Assisted_Selection_in_Beef_Cattle.pdf)

**Additional Readings:**

Alonso *et al.* (2006) Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in Arabidopsis. *Nat. Rev. Genet.* 1-12

Altmann T (1999). A complete BAC-based physical map of the Arabidopsis thaliana genome. *Nature Genetics* 1999 Jul; 22(3):271-5.

Borevitz *et al.* (2004) Genomics tools for QTL analysis and gene discovery. *Curr. Opin. Plant Biol.* 7, 132-136

Bortiriet *et al.* (2006) Advanced in maize genomics: the emergence of positional cloning. *Curr. Opin. Plant Biol.* 9. 164-171

Fujii J., Otsu K., Zorzato S. De Leon S., Khanna V.K., Weiler J.E., O'Brien P.J., MacLennan D.H. (1991). "Identification of a mutation in porcine ryanodine receptor associated with malignant hyperthermia". *Science* 253: 448.

Huang *et al.* (2005) Comparative genomics enabled the isolation of the *R3a* late blight resistance gene in potato. *Plant J.* 42, 251-261

Huang N., Cogburn L.A., Agarwal S.K., Marks H.L., Burnside, J. (1993). Over expression of a truncated growth hormone receptor in the sex-linked dwarf chicken: evidence for a splice mutation. *Molecular Endocrinology* 7: 1391.

Janderet *al.* (2002) Arabidopsis map-based cloning in the post-genome era. *Plant physiol.* 129, 440-450

Keller *et al.* (2005) Map-based isolation of disease resistance genes from bread wheat: cloning in a surprise genome. *Genet. Res.* 85, 93-100

King & Motulsky (2002). *Science* 298, 2342-2343 Rosenberg et al. (2002). *Science* 298, 2381-2385.

Klungland H., Vage D.I., Raya L.G., Adalsteinsson S., Lien S. (1995). "The role of melanocyte-stimulating hormone (MSH) receptor in bovine coat color determination". *Mammalian Genome* 6: 636.

Lukowitzet *al.* (2000) Positional cloning in *Arabidopsis*. Why it feels good to have a genome initiate working for you. *Plant Physiol.* 123, 795-805

Maria Sicaet *al.* (2005) ISSR markers show differentiationn among Italian populations of *Asparagus acutifolius* L, *BMNC Genetics* . 6:17

Martin *et al.* (1993) Map-based cloning of a protein *kinase* gene conferring disease resistance in tomato. *Science* 262, 1432-1436

Montgomery G.W., Lord E.A., Penty J.M., Dodds K.G., Broad T.E., Cambridge L.M., Sunden S.L.F., Stone R.T., Crawford A.M. (1994). "The Booroola fecundity (FecB) gene maps to sheep chromosome 6". *Genomics* 22 148.



Morton N.E., (1955). “Sequential tests for the detection of linkage”. *American journal of human genetics* 7: 277.

Peters *et al.* (2003) Forward genetics and map-based cloning approaches. *Trends Plant Sci.* 8, 484-491

Salviet *al.* (2005) to clone or not to clone plant QTLs: present and future challenges. *Trends Plant Sci.* 10. 297-304

Tai *et al.* (1999) Expression of *Bs2* pepper gene confers resistance to bacterial spot disease in tomato. *PNAS* 96, 14153-14158

Vaiman D., Koutita O., Oustry A., Elsen J.M., Manfredi E., Fellous M., Cribiu E.P. (1996). “Genetic mapping of the autosomal region involved in XX sex-reversal and horn development in goats”. *Mammalian Genome* 7:133.